



DEFINICE METADATOVÝCH FORMÁTŮ	31. října 2019 Dokument verze 0.1
Definice metadatových formátů pro digitalizaci zvukových dokumentů Část 2 Fonografické válečky	

Autoři:

Natalie Ostráková, Pavlína Kočišová, Miroslava Beňačková, Kateřina Šírová Vojířová, Filip Šír

Historie verzí

jméno	datum	verze dokumentu	provedené změny
Pavčina Kočiřová, Miroslava Beňáčková, Natalie Ostráková, Kateřina Vojířová Šírová, Filip Šír	říjen 2019	0.1	První draft

Tato definice metadatových formátů vychází z Definice metadatových formátů pro digitalizaci zvukových dokumentů ve verzi 0.3 v rámci Standardu Národní digitální knihovny (dále Standard NDK).

Obsah

Autoři:.....	1
1 Popis standardu	4
1.1 Účel standardu	4
1.2 Zodpovědnost	4
1.3 Dohoda mezi dodavatelem dat a zadavatelem	4
1.4 Popis a verze použitých standardů a identifikátorů	5
1.5 Definice pojmů	7
1.6 Terminologie	8
2 Výstupy digitalizace.....	8
3 Granularita metadatového záznamu	11
4 Identifikátory	11
5 Struktura SIP balíčku	12
5.1 Soubor info	14
5.2 Složka pro zvukové soubory sourceaudio	16
5.3 Složka pro zvukové soubory mastercopyaudio	17
5.4 Složka pro zvukové soubory usercopyaudio	17

5.5 Složka pro obrazové soubory mastercopy	17
5.6 Složka pro obrazové soubory usercopy	17
5.7 Složka alto	17
5.8 Složka txt	17
5.9 Složka amdsec	17
5.10 Složka pro katalogizační záznam catalog_entry	19
5.11 Soubor main_mets	19
5.12 Checksum	19
6 Názvová konvence složek a souborů	20
7 Metadata	24
7.1 Kořenový element hlavního METS záznamu	24
7.2 METS hlavička <metsHdr>	24
7.3 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core – katalogizace dle RDA pravidel	25
7.3.1. Pole MODS a DC pro úroveň zvukového dokumentu	28
7.3.2 Pole MODS a DC pro úroveň skladby	37
7.3.3 Pole MODS a DC pro obrazové informace (přílohy)	46
7.4 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core – katalogizace dle AACR2 pravidel	50
7.4.1. Pole MODS a DC pro úroveň zvukového dokumentu	53
7.4.2 Pole MODS a DC pro úroveň skladby	60
7.4.3 Pole MODS a DC pro obrazové informace (přílohy)	69
7.5 Technická a administrativní metadata ve vedlejších METS souborech	74
7.5.1 PREMIS Object	76
7.5.2 PREMIS Event	82
7.5.3 PREMIS Agent	85
7.5.4 Technická metadata pro zvukové soubory	88
7.5.5 Technická metadata MIX	110
7.6 METS část <fileSec>	118
7.6.1 <fileSec> hlavního záznamu METS	118
7.6.2 <fileSec> vedlejšího záznamu amd_mets.xml	121
7.7 METS část <structMap> - Strukturální metadata	122
7.7.1 <structMap> hlavního záznamu METS	122

7.7.2 <structMap> vedlejšího záznamu METS (AMD_METS.xml)	124
7.8 OCR (ALTO XML a TXT OCR)	125
8 Autorsko-právní metadata	142

1 Popis standardu

1.1 Účel standardu

Definice metadatových formátů pro zvukové dokumenty (dále DMF) bude sloužit jako předpis pro výsledek procesu digitalizace zvukových dokumentů v digitalizačních projektech knihoven v ČR. V první části byl standard zaměřen na gramofonové desky. Nyní se standard zaměřuje na fonografické válečky. Bude sloužit jako jednotný formát pro paměťové instituce, které chtějí svá data dlouhodobě archivovat v některém ze systémů dlouhodobé archivace využívajícím Standard NDK.

Pro koho není určen? DMF nebude určen pro malé lokální digitalizační aktivity, které mají jako jediný účel zpřístupnit dokumenty pro omezený okruh uživatelů. Ty si z něj mohou brát inspiraci, ale nemusí ho do důsledku dodržet, pokud nemají v plánu dlouhodobé uložení takto vytvořených dat.

V této specifikaci je blíže definován typ "zvukový dokument".

1.2 Zodpovědnost

Za DMF, jeho správnost a rozvoj, je zodpovědná Národní knihovna ČR. Kontaktním oddělením je Odbor digitálních fondů, resp. Oddělení pro standardy. S doplňujícími dotazy lze kontaktovat Pavlínu Kočišovou (pavlina.kocisova@nkp.cz, deskriptivní metadata) a Natálii Ostrákovou (natalie.ostrakova@nkp.cz, technická metadata). Standard vznikl ve spolupráci s Národním muzeem (kontaktní osoba Filip Šír, filip.sir@novyfonograf.cz).

1.3 Dohoda mezi dodavatelem dat a zadavatelem

DMF definuje podobu dat a metadat vyrobených dodavatelem a slouží jednak jako závazný dokument pro zadání digitalizace a jednak pro kontrolu dodaných dat. Požadavky definuje zadávající instituce a ošetřuje s dodavatelem dat smlouvou. Jedná se např. o:

- granularita záznamů - zadavatel si musí určit, v jakém rozsahu trvá na vyplnění nepovinných metadatových elementů. DMF obsahuje definici toho, co lze vytvořit, ale je na zadavateli, aby definoval požadovaný rozsah. Z hlediska dlouhodobého uchování jsou nezbytné všechny elementy s označením M a jim podřazené elementy.

- podrobnosti skenování obrazů - např. použité rozlišení (300 nebo více)

1.4 Popis a verze použitých standardů a identifikátorů

V této verzi specifikace budou použity následující verze standardů:

Standard	Verze	odkaz na webové stránky standardu
METS	verze 1.11	http://www.loc.gov/standards/mets/
MODS	verze 3.6	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
AES57	verze 1.0	http://www.aes.org/publications/standards/
PREMIS	verze 2.2	http://www.loc.gov/standards/premis/
ALTO	verze 3.1	http://www.loc.gov/standards/alto/
CopyrightMD	verze 0.91	http://www.cdlib.org/groups/rmg/

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

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

V této verzi specifikace mohou být použity následující identifikátory:

UUID	http://www.ietf.org/rfc/rfc4122.txt	
Objednací číslo	katalogové číslo nakladatele - nakladatelské číslo, které je přiřazeno jednomu válečku = issue number	1 váleček = 1-n nahrávek
URN:NBN	https://resolver.nkp.cz	Pro jednotlivé dokumenty se samostatným bibliografickým záznamem.
Matriční číslo	matrix number, unikátní pro konkrétní nahrávku	Vyskytuje se zřídka
Další identifikátor	upc, music publisher, local	Je možné použít další standardizovaný identifikátor, v případě nejasností kontaktujte Oddělení pro standardy.

Povinné předpoklady:

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

Další doporučení:

- původní neupravené zvukové soubory by měly mít parametry 96kHz, 24bit, zvukový formát WAV (specifikace: původní z r. 1991 [Multimedia Programming Interface and Data Specifications 1.0](#) s. 56-65 a update z r. 1994 [New](#)

[Multimedia Data Types and Data Techniques](#) s. 12-22) nebo formát BWF (specifikace: [Specification of the Broadcast Wave Format \(BWF\)-Version 1](#) a [Version 2](#))

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

Význam pole "Povinnost":

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

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

1.5 Definice pojmů

- **SA** = source audio = původní neupravený zvukový soubor
- **MC** = master copy = archivní kopie zvukového nebo obrazového záznamu (MCA, MC)
- **UC** = user copy = uživatelská kopie zvukového nebo obrazového souboru (UCA, UC)
- **PS** = původní sken – obrazový soubor vzniklý při digitalizaci, který se po zpracování (ořez, narovnání apod.) maže a dále se neukládá
- **základní intelektuální entita** = jedna zvuková kolekce (viz dále)
- **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 zvukovou kolekci

1.6 Terminologie

Bibliografický záznam = užíváno ve významu metadatového popisu obsahu zvukového dokumentu

Cover = obal zvukového nosiče; bude se skenovat zátka a obal válečku; případně i spodek obalu, pokud jsou na něm užitečné informace)

IMGdisc = sken zvukového nosiče (skenovat se bude okružní válečku a taky okružní a zátka válečku vedle sebe)

Katalogové číslo nakladatele = nakladatelské číslo = issue number, které je přiřazeno zvukovému dokumentu a které může být nalezeno na fonografickém válečku (okružní nosiče) nebo na zátce (obalu válečku)

Matriční číslo = matrix number, které je přiřazeno konkrétní nahrávce na fonografickém válečku; jeho výskyt není častý

Fonováleček = druh konkrétního zvukového nosiče

Skladba = zvuková nahrávka uceleného díla/celku, případně jeho ekvivalentu, za kterou může být autorský počin vytvoření, interpretace a záznamu díla

Zvuková kolekce = v případě fonografických válečků jde o soubor jednoho nebo více fonografických válečků v 1 obalu nebo více obalech (ve standardu je zaměňována za termín zvukový dokument)

Zvuková nahrávka = záznam zvuku na jakémkoliv nosiči včetně virtuálního

Zvuková stopa = podmnožina zvukové nahrávky, nosič může mít více stop-tracků

Zvukový dokument = pramen, jehož informačním obsahem je zaznamenaný/nahráný zvuk, je tvořen zvukovým nosičem nebo souborem nosičů (jedním či více fonografickými válečky)

Zvukový soubor = datový soubor se zvukovou nahrávkou nebo stopou, trackem

Zvukový záznam = v katalogizační terminologii: metadatový záznam o zvukovém dokumentu

2 Výstupy digitalizace

1. Původní neupravené zvukové soubory (WAV, 96kHz, 24bit) - **source_audio (SA)**

1.1. Jedna či více skladeb (soundrecording) nacházejících se na zvukovém dokumentu (na jednom či více fonografických válečkách).

- 1.2. Bez úprav, bez použití ekvalizačních křivek
 - 1.3. Parametry pro SA vycházejí z metodiky pro digitalizaci https://docs.google.com/document/d/15iRU-77WksR9jfHinH3s6jPhda_a9BwENnb16beRaQg/edit a z doporučení pro digitalizaci zvukových dat NDK dostupného zde: <http://www.ndk.cz/standardy-digitalizace/standardy-pro-zvukova-data>
 - 1.4. povinné
 - 1.5. ekvivalentní pojmy: flat audio, preservation master
2. Archivní kopie zvukových souborů (např. WAV) - **master_copy_audio (MCA)**
 - 2.1. Kopie odvozená ze source_audio
 - 2.2. Jedna či více skladeb (soundrecording) nacházejících se na zvukovém dokumentu (na jednom či více fonografických válečkách)
 - 2.3. nepovinné - v případě, že se nebude od SA ničím lišit, zachová se pouze SA, MCA je nepovinné (kvůli objemu dat)
 - 2.4. ekvivalentní pojmy: remaster, production master
3. Uživatelské kopie zvukových souborů (např. MP3) - **user_copy_audio (UCA)**
 - 3.1. Kopie odvozené z master_copy_audio (MCA), v případě, kdy není MCA, vychází UCA ze source_audio
 - 3.2. Jedna či více skladeb (soundrecording) nacházejících se na zvukovém dokumentu (na jednom či více fonografických válečkách)
 - 3.3. Parametry pro UCA vycházejí z doporučení pro digitalizaci zvukových dat NDK dostupné zde: <http://www.ndk.cz/standardy-digitalizace/standardy-pro-zvukova-data>
 - 3.4. povinné
 - 3.5. ekvivalentní pojmy: access copy
4. Archivní kopie obrazových souborů (JPEG 2000) - **master_copy (MC)**
 - 4.1. zátka, obal, spodek obalu, okruží, okruží a zátka vedle sebe
 - 4.2. doporučené rozlišení je 300 PPI a více, v plné barvě a ve formátu bezztrátový JPEG2000

5. Uživatelské kopie obrazových souborů (JPEG 2000) - **user_copy (UC)**

5.1. zátka, obal, spodek obalu, okruží, okruží a zátka vedle sebe

5.2. uživatelská kopie se ukládá ve ztrátovém formátu JPEG2000

6. Metadata pro audiosoubory

6.1. bibliografická metadata – MODS a DC

6.2. strukturální metadata – METS

6.3. technická metadata - PREMIS, AES57

METS nebude obsahovat technická metadata pro UC, jen pro SA a MC.

6.4. administrativní metadata – PREMIS, METS

7. Metadata obrazových souborů

7.1. bibliografická metadata – MODS a DC

7.2. strukturální metadata – METS

7.3. technická metadata – MIX, PREMIS

METS nebude obsahovat technická metadata pro UC, jen pro MC.

7.4. administrativní metadata – PREMIS, METS

8. OCR - ALTO XML_soubor pro každou stránku

Plnění bude dobrovolné

9. OCR TXT soubor - pro možnost stáhnout si jen text dokumentu (tam kde kvalita OCR je odpovídající), vyhledávání/indexace.

Plnění bude dobrovolné

10. kontrolní metadatové soubory ([s kontrolními součty](#) a [údaji o vzniku dat](#) apod.)

Pozn. odkazy vedou na kapitolu s příslušnou problematikou.

3 Granularita metadatového záznamu

Zvukový dokument obsahuje úrovně:

- zvukový dokument (soundcollection), který je tvořený zvukovým nosičem nebo jejich souborem. Obsahuje zvukovou nahrávku (tj. skladby nebo soubor skladeb), vydanou jako celek/komplet.
- skladba (soundrecording), tedy zvuková nahrávka uceleného díla/celku, případně jeho ekvivalentu.
- související obrazové informace:
 - o obal zvukového nosiče (cover – skenuje se zátka, obal válečku a případně spodek obalu
 - o zvukový nosič (imgdisc) – skenuje se okružní válečku a okružní a zátka vedle sebe
- Základní intelektuální entitou je zvukový dokument, reprezentovaná jedním strojově čitelným jedinečným identifikátorem (např. čárový kód, RFID, katalogové číslo (issue number) apod.)
- Všechny metadatové záznamy jsou zakotveny v METS formátu, 1 kolekce = 1 METS dokument
- SIP balíček je tvořen pro jeden zvukový dokument (tzn., zahrnuje všechny informace týkající se celku, válečků, jednotlivých stop a obrazových informací)

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 intelektuální entity
- syntax URN:NBN musí odpovídat specifikaci identifikátoru URN:NBN pro resolver NK (např. urn:nbn:cz:nk-12345a pro výstupy z projektu NDK)

Doporučené identifikátory jsou následující:

- Strojově čitelný identifikátor (např. RFID, čárový kód EAN apod.) - identifikuje celou zvukovou kolekci a celý zvukový dokument dle doporučení metodiky (jeden váleček)
- Katalogové číslo nakladatele (pole 028 ind1=0 \$a v MARC 21) - nakladatelské číslo = issue number, které je přiřazeno desce nebo kolekci válečků
- matriční číslo = matrix number (pole 028 ind1=1 \$a v MARC 21), které je přiřazeno konkrétní nahrávce
- UUID - nutné použít pro všechny úrovně digitálního objektu
- URN:NBN pro dokument, který má samostatný bibliografický záznam, bude se přidělovat během digitalizace. Povinně je přidělován na vrcholové úrovni popisu.
- Další standardizované identifikátory (music-publisher, upc...)

5 Struktura SIP balíčku

Souborová struktura SIP balíčku:

OBSAHUJE SLOŽKU/ SOUBOR	OBSAHUJE	POVINNOST	POJMENOVÁNÍ SLOŽKY
soubor info	popis obsahu balíčků (manifest)	ANO	

soubor main_mets	hlavní METS záznam	ANO	
složka source_audio	původní neupravené zvukové soubory	ANO	source_audio
složka master_copy_audio	zvukové soubory – archivní kopie	ANO/NE viz 2.3	master_copy_audio
složka user_copy_audio	zvukové soubory – uživatelské kopie	ANO	user_copy_audio
složka master_copy	obrazové soubory – archivní kopie	ANO	master_copy
složka user_copy	obrazové soubory – uživatelská kopie	ANO	user_copy
složka alto	soubory ocr alto pro každou stranu	NE	Alto
složka txt	soubory ocr text pro každou stranu	NE	Txt
složka amd_sec	soubory vedlejších mets pro všechny reprezentace jedné zvukové stopy či obrazové stránky dokumentu	ANO	amd_sec

složka catalog_entry	katalogizační záznam zvukového dokumentu ve formátu marcxml v samostatné složce	NE	catalog_entry
soubor checksum	kontrolní součet balíčku	ANO	

Technická a administrativní metadata nesmí být obsažena v hlavním METS záznamu. Každá stránka musí mít technická a administrativní metadata ve svém souboru s METS záznam (amd_mets.xml), stejně jako každý audio soubor.

SIP balíček = 1 adresář pro zvukový dokument. Platí i v případě, že jde o kolekci několika válečků společně popsaných.

Hlavní složka SIP balíčku obsahuje následující složky a soubory:

5.1 Soubor info

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.

element	atributy	popis	Povinnost
<info>		kořenový element INFO záznamu	M
<created>		časový údaj o vzniku balíčku ve formátu ISO8601 na úroveň vteřin	M

<metadata version>		verze metadatové specifikace, podle které byl balíček zpracován povolené verze: 0.1	M
<packageid>		název kořenového adresáře balíčku viz kap. 6	M
<mainmets>		název hlavního METS souboru včetně přípony	M
<validation>			MA
	version	verze validačního nástroje (při použití Komplexního validátoru stačí jen verze, v případě použití jiného validačního nástroje nutno vyplnit i jeho název)	R
		výstup validačního nástroje (např. OK; Valid)	R
<titleid>		soupis identifikátorů titulu - vypsát všechny možné (viz type), které titul má přidělen	M
	type	preferované hodnoty: urnnbn, uuid, issue number, matrix number; alespoň jedna z nich musí být nutně přítomna	M
<collection>		údaje o větším celku (projektu), pokud do některého balíček patří - např. digitalizace pro ANL	R
<institution>		název instituce, která je zadavatelem digitalizace	R

<creator>		tvůrce balíčku – kód instituce (firmy), která balíček vytvořila	M
<size>		velikost balíčku v kB - bez souboru info.xml	M
<itemlist>		obsahuje seznam všech souborů v balíčku vč. souborů v podadresářích a souboru info.xml	M
	itemtotal	celkový počet souborů	M
<item>		obsahuje cestu k jednomu souboru "txt\txt_nk-00027x_0001.txt" (takto definovaná cesta je doporučena, jsou možné i jiné podoby zápisu, které odpovídají logickým pravidlům, nejsou však doporučeny; cesta může být uvozena maximálně jednou tečkou). element je opakován podle nutnosti - tj. podle počtu souborů	M
<checksum>		odkaz na soubor s MD5 a jeho MD5 kontrolní součet	M
	type	bude vždy "md5"	M
	checksum	kontrolní součet souboru - použit stejný algoritmus jako byl použit v MD5 souboru	M
<note>		poznámka	O

5.2 Složka pro zvukové soubory sourceaudio

Složka s původními neupravenými zvukovými soubory (jedna strana až více stran válečků ve formátu WAV). Produkce zvukových souborů by se měla řídit Standardem NDK (<https://www.ndk.cz/standardy-digitalizace/standardy-pro-zvukova-data>).

5.3 Složka pro zvukové soubory mastercopyaudio

Složka s master kopiemi obsahuje archivní kopie všech zvukových souborů, např. ve formátu WAV. 1 soubor = 1 stopa (jedna strana až více stran válečků ve formátu WAV). Bude se uchovávat pouze jeden druh master kopií i přesto, že může vzniknout více druhů master kopií v případě různých druhů úprav. V případě, že se nebude MC nijak lišit od SA, potom tato složka není povinná.

5.4 Složka pro zvukové soubory usercopyaudio

Složka s uživatelskými kopiemi, všechny uživatelské kopie např. ve formátu MP3. 1 soubor = 1 skladba

5.5 Složka pro obrazové soubory mastercopy

Složka s archivními kopiemi všech obrazových souborů např. v bezztrátovém JPEG2000 (obal válečků, obrazy fyzických nosičů). Produkce obrazových dat by se měla řídit doporučenými standardy pro digitalizaci (<http://www.ndk.cz/standardy-digitalizace/standardy-pro-obrazova-data>).

5.6 Složka pro obrazové soubory usercopy

Složka s uživatelskými kopiemi všech obrazových souborů.

5.7 Složka alto

Obsahuje ke každé obrazové stránce 1 ALTO XML soubor, tj. tolik ALTO XML souborů, kolik je stránek patřících k jedné kolekci. Plnění je volitelné.

5.8 Složka txt

Obsahuje ke každé obrazové stránce 1 OCR soubor jako čistý text, tj. tolik OCR.TXT souborů, kolik je stránek patřících k jedné kolekci. Plnění je volitelné.

5.9 Složka amdsec

Složka s technickými metadaty musí obsahovat pro každý zvukový soubor i každou naskenovanou stránku 1 METS soubor amd_mets.xml. Z názvu souboru je patrné k jakému archivovanému souboru metadata patří. METS soubory popisující zvuková data obsahují v názvu „sa“ a „mca“, tj. amd_mets_sa.....xml a amd_mets_mca.....xml, obrazová data obsahují pouze amd_mets.....xml. Každý soubor amd_mets.xml obsahuje následující části METS formátu:

- **amdSec** - administrativní metadata - obsahuje části:

- technická metadata **techMD**, která:
 - ve formátu **PREMIS Object** popisují vlastnosti:
 - původního zvukového souboru (SA), ze kterého vznikly archivní zvukové kopie, např. soubory WAV
 - původního obrazového souboru (PS), ze kterého vznikly archivní obrazové kopie, např. soubory TIFF
 - archivní kopie zvukového souboru (není povinné v případě, že se MC nijak neliší od SA)
 - archivní kopie obrazového souboru
 - ALTO XML, pokud se tvoří
 - případné jiné reprezentace - obraz i zvuk v případě opravy
 - ve formátu **AES57** popisují vlastnosti:
 - původního zvukového souboru
 - archivní kopie zvukového souboru (není povinné v případě, že se MC nijak neliší od SA)
 - ve formátu **MIX** popisují vlastnosti:
 - archivní kopie obrazu (veškeré generace archivních kopií)
 - 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 jednu stránku, který vzniká primárně k zachycení technických a administrativních metadat, bude odkazovat na soubory souvisejícími s tou konkrétní stránkou, tj. archivní kopie zvuku, původní zvukové soubory nebo archivní kopie obrazu, a ALTO XML a OCR TXT, pokud existují
- **structMap** - pouze fyzická strukturální mapa, povinná část METS záznamu, bude ukazovat strukturu souborů k dané stopě nebo stránce, tj. source audio, master copy zvuku a obrazu, a ALTO XML a OCR TXT, pokud existují
- **copyrightMD** - autorsko-právní Metadata je možné definovat i pro jednotlivé soubory. Možnosti a pravidla dědičnosti jsou uvedeny v příslušné kapitole (kapitola 8 – Autorsko-právní Metadata).

5.10 Složka pro katalogizační záznam `catalog_entry`

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

5.11 Soubor `main_mets`

Hlavní METS záznam bude obsahovat části:

- `dmdSec` – bibliografická metadata ke každé úrovni zvukového dokumentu ve formátu MODS a DC, včetně obrazových informací
- `amdSec` - autorsko-právní metadata k zvukovému dokumentu 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 soubor (fyzická i logická část)
- `structLink` - vazby mezi fyzickou a logickou částí strukturální mapy (tedy mezi jednotlivými zvukovými stopami a stranou válečku a mezi obrázky a válečkem či kolekcí)

5.12 Checksum

Pro kontrolní součty je využíván algoritmus MD5. Soubor s kontrolním součtem souborů pro 1 SIP balíček. Také obsahuje kontrolní součet pro každý soubor obsažený v SIP balíčku (kromě `info.xml` a `*.md5` souboru samotného).

Obsah souboru je definován následující gramatikou ve formátu ABNF¹

SOUBOR = *RADEK

RADEK = HODNOTA-MD5 MEZERA NAZEV-SOUBORU KONEC-RADKU

HODNOTA-MD5 = 32HEXDIG

MEZERA = " " / TAB

NAZEV-SOUBORU = *SEGMENT

¹ <http://tools.ietf.org/html/rfc4234>

KONEC-RADKU = (CR LF) / LF
 SEGMENT = PATH-SEP FILENAME-CHARS
 PATH-SEP = "/" / "\"
 TAB = %x09
 CR = %x0D
 LF = %x0A
 FILENAME-CHAR = ALPHA / DIGIT / "." / "_" / "-"
 FILENAME-CHARS = 1*FILENAME-CHAR

kde sémantika je následující:

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

6 Názvová konvence složek a souborů

V případě použití např. čárového kódu jako identifikátoru bude kód sloužit k jednoznačné identifikaci jednoho zvukového dokumentu. Bude také tvořit základ pojmenování souborů. Stejný postup platí i v případě použití jiných strojově čitelných identifikátorů.

Pojmenování SIP balíčku

- každý SIP balíček přicházející z digitalizace musí obsahovat pouze jedinou intelektuální entitu (zvukový dokument o počtu jeden a více válečků).
- název balíčku **musí vycházet** z identifikátoru této entity, tj. URN:NBN, UUID nebo čárového kódu

1. V případě URN:NBN se do názvu balíčku převezme druhá část identifikátoru za urn:nbn:cz (např. z urn:nbn:cz:osa001-001kl vznikne název balíčku osa001-001kl osa001-001kl).
 2. V případě převzetí názvu z UUID se převezme celá část identifikátoru za prefixem uuid (tj. z uuid:21d5eff0-d9aa-11de-a7ba-000d606f5dc6 vznikne název balíku 21d5eff0-d9aa-11de-a7ba-000d606f5dc6)
- každý zvukový dokument 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 a pomlčka

Pojmenování složek

- viz návrh struktury SIP balíčku (master_copy_audio, user_copy_audio atp.)
- názvy nesmí obsahovat mezery, diakritiku a dvojtečku, doporučené oddělovače jsou podtržítka a pomlčka.

Pojmenování souborů

V případě použití uuid:

- o zvukové soubory nesou označení ve formě „druh-souboru_uuid_XXXX.wav, kde se druhem souboru rozumí „sa“, „mca“ nebo „uca“, XXXX je pořadové číslo skladby, př.:
 - o sa_b50eb6b0-f0a4-11e3-b72e-005056827e52_0001.wav
 - o mca_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.wav
 - o uca_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.mp3
- o **obraz válečku** - 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
 - o mc_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.jp2
 - o uc_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.jp2
- o **obraz obalu válečku** - skeny označené jako "druh-souboru_uuid_XXXX.jp2", kde druh souboru je "mc" nebo "uc", XXXX je pořadové číslo skenu; stejné pojmenování ponese i uživatelská kopie
 - o mc_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.jp2

- o uc_b50eb6b0-f0a4-11e3-b72e-005056827e52_0002.jp2

- o **hlavní mets** – bude mít označení main_mets_uuid.xml
 - o main_mets_b50eb6b0-f0a4-11e3-b72e-005056827e52.xml

- o **soubor info** - bude mít označení info_uuid.xml
 - o info_123456789.xml

- o **soubor ALTO** - bude mít označení “alto_uuid_XXXX.xml, kde “alto” určuje druh souboru a XXXX pořadové číslo xml souboru
 - o alto_b50eb6b0-f0a4-11e3-b72e-005056827e52_0001.xml

- o **soubor OCR TXT** - označení “txt_uuid_XXXX.txt, kde “txt” určuje druh souboru a XXXX pořadové číslo txt souboru
 - o txt_b50eb6b0-f0a4-11e3-b72e-005056827e52_0001.txt

- o **soubor MD5** - označení “md5_uuid.md5”, kde “md5” určuje druh souboru a XXXX určuje pořadové číslo md5 souboru
 - o md5_b50eb6b0-f0a4-11e3-b72e-005056827e52.md5

- o **soubor calog_entry** – bude mít označení „cat_entry_uuid.xml
 - o cat_entry_b50eb6b0-f0a4-11e3-b72e-005056827e52_0001.xml

V případě použití čárového kódu:

- o zvukové soubory - pojmenované jako "druh-souboru_čárový-kód_XXXX.wav", kde druh souboru je “sa”, “mca” nebo “uca”, XXXX je pořadové číslo skladby, př.:
 - o sa_1234567890_0001.wav, sa_1234567890_0002.wav
 - o mca_1234567890_0001.wav, mca_1234567890_0002.wav
 - o uca_1234567890_0001.wav, uca_1234567890_0002.wav
 - o anebo v případě rozdělení na stopy: uca_1234567890_0001_0001.mp3,
uca_1234567890_0001_0002.mp3, uca_1234567890_0001_0003.mp3,
uca_123456789_0001_0004.mp3, uca_1234567890_0002_0001.mp3,

uca_1234567890_0002_0002.mp3,
uca_123456789_0002_0004.mp3

uca_1234567890_0002_0003mp3,

- o **obraz válečku** - skeny označené jako "druh-souboru_čárový_kód_XXXX.jp2", kde druh souboru je "mc" nebo "uc" a XXXX je pořadové číslo skenu; stejné pojmenování ponese i uživatelská kopie
 - o mc_1234567890_0001.jp2
 - o mc_1234567890_0002.jp2
 - o uc_1234567890_0001.jp2
 - o uc_1234567890_0002.jp2
- o **obraz obalu válečku** - skeny označené jako "druh-souboru_čárový_kód_XXXX.jp2", kde druh souboru je "mc" nebo "uc", XXXX je pořadové číslo skenu; stejné pojmenování ponese i uživatelská kopie
 - o mc_1234567890_0003.jp2
 - o uc_1234567890_0003.jp2
- o **hlavní mets** – bude mít označení main_mets_čárový_kód.xml
 - o main_mets_123456789.xml
- o **soubor info** - bude mít označení info_čárový_kód.xml
 - info_123456789.xml
- o **soubor ALTO** - bude mít označení "alto_čárový_kód_XXXX.xml, kde "alto" určuje druh souboru a XXXX pořadové číslo xml souboru
 - o alto_1234567890_0001.xml
- o **soubor OCR TXT** - označení "txt_čárový_kód_XXXX.txt, kde "txt" určuje druh souboru a XXXX pořadové číslo txt souboru
 - o txt_1234567890_0001.txt
- o **soubor MD5** - označení "md5_čárový_kód.md5", kde "md5" určuje druh souboru a XXXX určuje pořadové číslo md5 souboru
 - o md5_1234567890_0001.md5
- o **soubor calog_entry** – bude mít označení „cat_entry_čárový_kód.xml
 - o cat_entry_1234567890.xml

Názvy jakýchkoliv souborů náležející k jedné intelektuální entitě musí být založeny na jednom typu identifikátoru. Názvy musí být pouze malými písmeny a názvy nesmí obsahovat mezery ani diakritiku, doporučené oddělovače jsou podtržítka a pomlčka.

7 Metadata

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

7.1 Kořenový element hlavního METS záznamu

Kořenový element hlavního METS záznamu 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 zvukové kolekce, včetně roku vydání (název, rok vydání)	M
	TYPE	Hodnota vždy „audio cylinder“	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 (na úroveň vteřin)	M
	CREATEDATE	Datum vytvoření záznamu, musí být ve tvaru ISO 8601 (na úroveň vteřin)	M
<agent>		Údaje o tvůrci záznamu METS	M
	ROLE	Hodnota „CREATOR“	M

	TYPE	Hodnota „ORGANIZATION“	M
<name>		Jméno jednotlivce nebo organizace; tvůrce záznamu, buď dodavatel (firmaXY) nebo v případě tvorby záznamu v knihovně bude využita sigla knihoven, tj. pro NKČR hodnota "ABA001"	M
<agent>		Údaje o vlastníkovi METS	M
	ROLE	Hodnota „ARCHIVIST“	M
	TYPE	Hodnota „ORGANIZATION“	M
<name>		Jméno organizace; vlastník záznamu, v případě tvorby záznamu v knihovně bude vždy využita sigla knihoven, tj. pro NK ČR hodnota „ABA001“	M

7.3 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core – katalogizace dle RDA pravidel

- 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 <http://www.loc.gov/standards/mods/mods-conversions.html>
- pro vytvoření DC z MODS formátu musí být použito (a dle potřeb knihovny upraveno) oficiální mapování Kongresové knihovny – viz <http://www.loc.gov/standards/mods/mods-conversions.html>
- U digitalizovaných dokumentů je bibliografický popis vytvářen primárně z pohledu popisu fyzické předlohy, nejde o popis elektronického dokumentu

Zvukový dokument – fonografický váleček

- základní intelektuální entitou pro popis je zvukový dokument, tj. v jednom METS záznamu, který bude obsahovat metadata a strukturu jednoho a více válečků, budou MODS záznamy k této kolekci

- zvukovým dokumentem se rozumí vrcholná úroveň popisu fonografického válečku, ať již se týká válečku jen s jednou stranou nebo souboru mnoha válečků. Zvukový dokument odpovídá bibliografickému záznamu v katalogu. metadata budou popisovat entity²:
 - **zvukový dokument (soundcollection)** – zvukový nosič nebo soubor zvukových nosičů, na kterém byla zvuková nahrávka (tj. skladby nebo soubor skladeb) vydávaný jako celek/ komplet. Může obsahovat jeden či více válečků v jednom obalu nebo ve více obalech, které jsou popsány jedním bibliografickým záznamem v katalogu a jejichž obsah spolu souvisí nebo na sebe navazuje. Zvukový dokument je nadřazená entita pro všechny následující.
 - **Skladba (soundrecording)** - popisuje zvlášť každou zvukovou nahrávku, tvořící ucelené dílo. ● **příloha (supplement)** - přílohou se rozumí volně vložená entita k intelektuální entitě, např. volně vložený booklet, obsah, reklamní či edukativní příloha apod.

Rozlišujeme dva druhy příloh:

- 1) **obal (cover)** - popis pro fyzický obal kolekce. Skenuje se zátka, obal válečku, a pokud je to přínosné, tak i spodek obalu válečku. Pomocí elementu <relatedItem> se odkazuje ke kolekci.
- 2) **obraz válečku (imgdisc)** - popis pro sken válečku. Skenuje se okruží válečku a taky okruží a zátka vedle sebe.

Přílohy se pomocí elementu <relatedItem> odkazují k nejvyšší popisované úrovni, tj. zvukovému dokumentu. 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 (např. soundcollection, soundrecording, supplement). V případě opakování částí stejného typu se bude opakovat tomu odpovídající počet MODS záznamů v jednom SIP balíčku.

² Toto pořadí nevyjadřuje logickou strukturu dokumentu, ta je popsána dále v DMF.

- 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.

Element	Atributy	Popis	Povinnost
<dmdSec>		Identifikátor dmdSec části METS záznamu	M
	ID	Pro <dmdSec> s popisem zvukového dokumentu hodnota "MODSMD_SOUND COLLECTION" a "DCMD_SOUND COLLECTION"	M
		Pro <dmdSec> s popisem skladby hodnota "MODSMD_SOUNDRECORDING" a "DCMD_SOUNDRECORDING"	M
		Pro <dmdSec> s popisem přílohy hodnota "MODSMD_SUPPL" a "DCMD_SUPPL"	M
<mdWrap>		Element obsahující vložené záznamy MODS	M
	MDTYPE	Hodnota „MODS“ pro záznamy v MODS, hodnota „DC“ pro záznamy v Dublin Core	M
	MDTYPEVERSION	Číslo verze MODS, hodnota pro záznamy v MODS, např. „3.6“	O
	MIMETYPE	Hodnota „text/xml“	O
<xmlData>			M

Pole MODS a Dublin Core pro jednotlivé úrovně zvukových dokumentů

- ID u elementu <mods>: Identifikátory budou začínat prefixy MODS_SOUND COLLECTION, MODS_SOUNDRECORDING, 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:

1) zvukový dokument :

- MODS_SOUNDCOLLECTION_0001
- DC_SOUNDCOLLECTION_0001

2) skladba:

- MODS_SOUNDRECORDING_0001
- DC_SOUNDRECORDING_0001

3) přílohy (skeny obalu a okruží):

- MODS_SUPPL_XXXX, kde XXXX je pořadové číslo přílohy, např. MODS_SUPPL_0001 je první příloha atd.
- DC_SUPPL_XXXX, kde XXXX je pořadové číslo přílohy, např. MODS_SUPPL_0001 je první příloha atd.

7.3.1. Pole MODS a DC pro úroveň zvukového dokumentu

Obsah sloupce "Popis":

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

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. pro úroveň zvukového dokumentu	M	

		„MODS_SOUNDCOLLECTIO N_0001“		
<titleInfo>		Název kolekce, souborný název (pro plnění lze použít katalogizační záznam), samozřejmě lze využít všech prvků a elementů MODS, které názvové informace popisují) Opakovatelný element	M	
	type	hlavní název bez type-pole 245 \$a Může existovat pouze jeden hlavní název bez type. type: alternative-pole 246 translated -pole 242 uniform -pole 130, resp. 240 abbreviated – pole 210	MA	
<title>		Názvová informace	M	<dc:title>
<subTitle>		Podnázev	MA	<dc:title>
<partNumber>		Číslo části zvukového dokumentu, pokud je dělen na části. Údaj se čerpá z polí 240 a 245, podpole \$n.	MA	<dc:title>
<partName>		Název části z pole 240 nebo 245, podpole \$p.	MA	<dc:title>
<nonSort>		Část názvu, která má být vynechána při vyhledávání (např. <nonSort>The </nonSort>)	O	
<name>		Údaje o odpovědnosti. Přebíhá se z 1XX a 7XX marc21. Opakovatelný element pro více autorů/různé role	MA	
	type	Použít jednu z hodnot: -personal -corporate -conference -family	MA	
<namePart>		Údaje o křestním jméně a příjmení apod. Pokud je to	MA	<dc:creator>

		možné, tak vyjádřit pro křestní jméno i příjmení. Více křestních jmen se zapíše do jednoho elementu <namePart>. Pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart>. Pokud známe datum narození a úmrtí autora, vyplnit ve tvaru RRRR-RRRR s atributem type="date".		Nutno do jednoho pole DC spojit jméno i příjmení
	type	Použít jednu z hodnot: -date (MA) -family (MA) -given (MA) -termsOfAddress (R) Pokud se nejedná o osobu, atribut type se nepoužije.		
<nameIdentifier>		Číslo národní autority	RA	
<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. Příklad: <mods:name> <etal>akol.</etal> </mods:name>	O	

<role>		Specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		Popis role – nutno použít kontrolovaný slovník, např. http://www.loc.gov/marc/relators/relaterm.html	M	
	type	Hodnota „code“	M	
	authority	Údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority=“marcrelator“	M	
<typeOfResource>		Hodnota „sound recording“, „sound recording–musical“ nebo „sound recording–non musical“ Mělo by se vyčítat z MARC21 katalogizačního záznamu, z pozice 06 návěští – hodnoty i, j	R	<dc:type>
<genre>		Bližší údaje a typu dokumentu Stanovená hodnota „sound recording“	M	<dc:type>model:soundrecording</dc:type>
<originInfo>		Informace u původu předlohy	M	
	eventType	Hodnoty dle druhého indikátoru pole 264: 264_1 „publication“ Hodnota 1 se uvádí, jestliže pole obsahuje údaje o nakladateli zdroje 264_2 „distributor“ Hodnota 2 se uvádí, jestliže pole obsahuje údaje o distributorovi zdroje Pole s dalšími indikátory se využívat nemusí		
<place>		Údaje o místě spojeném s vydáním, výrobou nebo	MA	<dc:coverage>

		původem popisovaného dokumentu		
<placeTerm>		Konkrétní určení místa vydání, např. Praha, resp. „xr“ pro ČR Odpovídá hodnotám z katalogizačního záznamu, pole 264 \$a, resp. pole 008/15-17	MA	<dc:coverage>
	type	Hodnota „code“ pro údaj z pole 008 Hodnota „text“ pro údaj z pole 264	M	
	authority	Hodnota „marccountry“ jen u údaje z pole 008	MA	
<publisher>		Jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala, odpovídá poli 264 \$b v MARC21 Pokud má titul více vydavatelů, přebírají se ze záznamu všichni (jsou v jednom poli 264)	MA	<dc:publisher>
<dateIssued>		Datum vydání předlohy, odpovídá hodnotě z katalogizačního záznamu, pole 264_1 \$c a poli 008/07-10. Pro všechny ostatní výskyty v poli 264 \$c využít element <dateOther>.	M	<dc:date>
	encoding	Hodnota „marc“ jen u údaje z pole 008	R	
	point	Hodnoty „start“ resp. „end“ jen u údaje pro rozmezí dat z pole 008	MA	
	qualifier	Možnosti dalšího upřesnění, hodnota „approximate“ pro data, kde nevíme přesný údaj	R	
<dateOther>		datum distribuce, předlohy tento element se využije v případě výskytu \$c v: 264_2 "distribution"	R	<dc:date>

	type	Hodnota „distribution“ v případě 264_2	M	
<edition>		Údaje o pořadí vydání, odpovídá poli 250 \$a katalogizačního záznamu	R	
<issuance>		Údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07 Hodnota „monographic“	M	
<physicalDescription>		Obsahuje údaje o fyzickém popisu zdroje předlohy (pole 007, 300, 306, 337, 338)	M	
<form>		Údaje o fyzické podobě dokumentu Hodnota „sound recording“ z údaje z 007/00 code s Hodnota „cylinder“ u údaje z 007/01 code e údaje o typu média a typu nosiče zdroje/předlohy: Hodnota „audio“ u údaje z 337 – viz kontrolovaný slovník www.loc.gov/standards/valuelist/rdamedia.html hodnota „audio cylinder“ u údaje z 338 – viz kontrolovaný slovník http://www.loc.gov/standards/valuelist/rdacarrier.html	M	<dc:format>
	authority	Hodnota „marccategory“ u údaje z 007/00 Hodnota „marcsmd“ u údaje z 007/01 Pole 337: „rdamedia“ Pole 338: „rdacarrier“	MA	

	type	Pole 337: „media“ Pole 338: „carrier“	MA	
<extent>		Údaje o rozsahu, odpovídá hodnotám v poli 300 \$a, \$b, \$c, \$e; 306 Příklad: 300 \$a 1 fonováleček 300 \$b analog, 160ot/m, mono 300 \$c průměr 54 mm, délka 106 mm	M	<dc:coverage>
<note>		Slouží k vepisování údajů z polí 5XX, 245 \$c. Pro každou poznámku musí být samostatný <note> element.	RA	<dc:description>
	type	Pro obsah pole 245 hodnota „statement of responsibility“ Pro obsah pole 508 hodnota „creation/production details“ Pro obsah pole 511 hodnota „performers“ Pro obsah pole 518 hodnota „venue“ Pro obsah pole 546 hodnota „language“ Pro obsah pole 561 hodnota „ownership“ Pro obsah pole 562 hodnota „version identification“ Pro obsah pole 583 hodnota „action“	MA	
<tableOfContents>		slouží k vepsání názvů skladeb na válečku; obsah 505\$a	MA	<dc:description>
	displayLabel	ind1=0 hodnota „Contents“ ind1=1 hodnota „Incomplete contents“ ind1=2 hodnota „Partial contents“	R	

<relatedItem>		<p>Informace o dalších dokumentech, jejich částech či zdrojích, které jsou ve vztahu k popisovanému zvukovému dokumentu.</p> <p>Dále lze použít pro vyjádření edice, ve které je dokument vydán (tzn. kolekce je součástí nějaké edice), údaj o edici musí obsahovat minimálně element <title> s jejím názvem.</p> <p>Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsány pro tyto elementy</p>	RA	
	type	např. hodnota „series“	MA	
<identifier>		Údaje o identifikátorech, obsahuje unikátní identifikátory nebo lokální	M	<dc:identifier>
	type	<p>Budou se vyplňovat následující hodnoty, pokud existují:</p> <ul style="list-style-type: none"> - issue number (MA) - uuid (M) - urn:nbn (M) - upc (MA) - local (MA) - music-publisher (MA) 	M	
<location>		Údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
<physicalLocation>		Údaje o instituci, kde je fyzicky uložen popisovaný dokument, např. MZK; nutno použít kontrolovaný slovník – sigly knihoven (BOA001); odpovídá poli 040 v MARC21	MA	<dc:source>

		Pozn. U dokumentů ze soukromých sbírek není možné vyplnit		
	authority	Hodnota „siglaADR“	M	
<shelfLocator>		Signatura nebo lokační údaje o dokumentu, který slouží jako předloha	M	<dc:source>
<recordInfo>		Údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		Popis standardu, ve kterém je přebíraný katalogizační záznam 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	MA	
<recordContentSource>		Kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	Hodnota „marcorg“	R	
<recordCreationDate>		Datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	Hodnota „iso8601“	M	
<recordChangeDate>		Datum změny záznamu, na úroveň minut podle normy ISO 8601	R	
	encoding	Hodnota „iso8601“	M	
<recordIdentifier>		Identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	Hodnota se přebírá z katalogu pole 003	R	
<recordOrigin>		Údaje o vzniku záznamu; hodnoty „machine generated“ nebo „human prepared“	R	
<languageOfCataloging>		Jazyk katalogového záznamu	R	

<languageTerm>		Přebírá se z katalogu – pole 040 \$b	R	
	authority	Hodnota „iso639-2b“	R	

7.3.2 Pole MODS a DC pro úroveň skladby

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	obsahuje název úrovně, pro úroveň skladby je to "MODS_SOUNDRECORDING_0001", kde číslo je pořadové číslo skladby v rámci zvukového dokumentu.	M	
<titleInfo>		názvové informace - pokud nemá skladba zjistitelný název, kopíruje se informace z nadřazené struktury	M	
	type	alternative uniform translated	MA	
<title>		hlavní název	M	<dc:title>
<subTitle>		podnázev	MA	<dc:title>
<partNumber>		číslo skladby či její pořadí	MA	<dc:title>

<partName>		Název části z pole 240 nebo 245, podpole \$p.	MA	<dc:title>
<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 corporate conference family	MA	
<namePart>		Údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení. Více křestních jmen se zapíše do jednoho elementu <namePart>. Pokud nelze rozlišit křestní jméno a příjmení, nepoužije se type a jméno se zaznamená v podobě, jaké je, do jednoho elementu <namePart>. Pokud známe datum narození a úmrtí autora, vyplnit ve tvaru RRRR-RRRR s atributem type="date".	MA	<dc:creator>
	type	Použít jednu z hodnot: date (MA) family (MA) given (MA) termsOfAddress (R) Pokud se nejedná o osobu, atribut type se nepoužije.	MA	

<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
	type	hodnota "code" – kód role z kontrolovaného slovníku (http://www.loc.gov/marc/relators/relaterm.html)	M	
	authority	údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"	M	
<genre>	type	stanovená hodnota „sound recording“	M	<dc:type>m odel:soundr ecording</d c:type>
<originInfo>		informace o původu předlohy: odpovídá poli 264	M	
	event Type	Hodnoty dle druhého indikátoru pole 264: 264_1 „publication“ Hodnota 1 se uvádí, jestliže pole obsahuje údaje o nakladateli zdroje 264_2 „distribution“ Hodnota 2 se uvádí, jestliže pole obsahuje údaje o distributorovi zdroj		

<place>		údaje o místě spojeném s vytvořením, vydáním, distribucí nebo výrobou popisovaného dokumentu odpovídá hodnotě 264 \$a	MA	<dc:coverage>
<placeTerm>		konkrétní určení místa vydání, např. Praha, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 264 \$a, resp. pole 008/15-17	MA	<dc:coverage>
	authority	authority - hodnota "marccountry" jen u údaje z pole 008	MA	
<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07 hodnota "monographic"	M	

<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy (pole 300,306)	M	
<extent>		údaje o rozsahu; odpovídá hodnotám v poli 300, 306 \$a	M	<dc:coverage>
<language>		údaje o jazyce dokumentu v případě vícenásobného výskytu nutno element <language> opakovat	M	

	objectPart	<p>možnost vyjádřit jazyk konkrétní části díla;</p> <p>možné hodnoty:</p> <ul style="list-style-type: none"> - sung or spoken text, odpovídá poli 041 \$d - translation, odpovídá poli 041 \$h - libretto, odpovídá poli 041 \$e -accompanying materiál, odpovídá poli 041 \$g 	MA	
<languageTerm >		<p>přesné určení jazyka – kódem; nutno použít kontrolovaný slovník ISO 639-2, http://www.loc.gov/standards/iso639-2/php/code_list.php</p>	M	<dc:language >
	type	hodnota „code“	M	
	authority	hodnota „iso 639-2b“	M	
<typeOfResource>		<p>Obsahuje hodnotu z návěstí z pozice 05:</p> <ul style="list-style-type: none"> - „sound recording-musical“ – odpovídá \$j - „sound recording-nonmusical“ – odpovídá \$i 	R	
<note>		<p>Obecná poznámka, může též sloužit jako element pro vyjádření údaje o odpovědnosti, poznámka o realizátorech, účinkujících či jazyku.</p> <p>Pro každou poznámku je potřeba samostatný element <note></p>	RA	<dc:description>

	type	<ul style="list-style-type: none"> - statement of responsibility“ – odpovídá poli 245 \$c - „creation/ production credits“ – odpovídá poli 508 \$a - „performers“ – odpovídá poli 511 \$a - „venue“ – odpovídá poli 518 \$a - „language“ – odpovídá poli 546 \$a - version identification – odpovídá poli 562 \$b - action – odpovídá poli 583 \$a 	M	
<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“.	M	
<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>

<geographic>		geografické věcné třídění použit 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>
<temporal>		chronologické věcné třídění použit 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>
<name>		jméno použité jako věcné záhlaví použit 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; struktura a atributy stejné jako pro údaje o původcích – viz element <name>	R	<dc:subject>
<title>		Pole použité jako vedlejší věcné záhlaví	R	<dc:subject>
<titleInfo>		Unifikovaný název	M	
<genre>		Použit' při 653 a ind2=6	R	<dc:subject>
<geographicCode>		Kód geografické oblasti, obsah pole 043	R	<dc:subject>

	authority	Hodnota „marcgac“	M	
<relatedItem>		<p>Informace o zvukovém dokumentu, jehož je skladba součástí</p> <p>Údaje o zvukovém dokumentu musí obsahovat nejméně element <title> s jeho názvem.</p> <p>Poznámka: element <relatedItem>. Může obsahovat jakýkoliv jiný elementn MODS – jejich použití se řídí pravidly popsanými pro tyto elementy.</p>	R A	
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální	M	<dc:identifier >
	type	<p>budou se povinně vyplňovat následující hodnoty, pokud existují:</p> <ul style="list-style-type: none"> - uuid (M) - „issue number“ – číslo válečku, v MARC21 pole 028 ind1=0 \$a, např. \$a DV 6196, přebírá se z katalogu, ind1=0 (MA) - type=“matrix-number“ pro unikátní matriční číslo konkrétní nahrávky, pole 028 ind1=1 \$a (MA) - music-publisher 028a, b ind1=3 (R) - upc, ean (R) - jiný interní identifikátor - hodnota atributu „local“ (R) 	M	

<recordInfo>		údaje o metadatovém záznamu válečku – jeho vzniku, změnách apod.	M	
<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam 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 A	
<recordContentSource >		kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"	M	
<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 na úroveň minut	R	
	encoding	hodnota atributu "iso8601"	M	

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

7.3.3 Pole MODS a DC pro obrazové informace (přílohy)

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	
<titleInfo>		názvové informace o příloze	M	
<title>		uazapíše se, která část fonoválečku je na skenu; hodnoty: „Zátka“, „Obal“, „Spodek obalu“, „Okružní“, „Zátka a okružní“	MA	<dc:title>

<typeOfResource>		Možné hodnoty „still image“	M	
<genre>		Hodnota „supplement“	M	<dc:type>model:supplement</dc:type>
	type	Na určení typu přílohy Povolené hodnoty: cover, imgdisc	M	
<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy	M	
<form>		Údaje o fyzické podobě dokumentu Hodnoty se přejímají z kontrolovaných slovníků		
	authority	Hodnota „marccategory“ Pole 337: authority=“rdamedia“ Pole 338: authority=“rdacarrier“	MA	<dc:format>
	type	Pole 337: „media“ Pole 338: „carrier“	MA	
<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	

<note>		obecná poznámka k příloze	O	<dc:description>
<relatedItem>		Popis zvukového dokumentu, ke kterému příloha patří. Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsány pro tyto elementy	M	
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální – viz přehled typů atributů níže	M	<dc:identifier>
	type	budou se povinně vyplňovat následující hodnoty, pokud existují: - uuid (M)	M	
<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam 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	MA	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila; nutno vytvořit kontrolovaný slovník	R	

<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"	M	
<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 na úroveň minut	R	
	encoding	hodnota atributu "iso8601"	M	
<recordOrigin>		údaje o vzniku záznamu; hodnoty „machine generated“ nebo „human prepared“	R	
<languageOfCataloging>		jazyk katalogového záznamu	R	
<languageTerm>		přebírá se z katalogu - pole 40 \$b	R	
	authority	hodnota "iso639-2b"	R	

7.4 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core – katalogizace dle AACR2 pravidel

- 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 <http://www.loc.gov/standards/mods/mods-conversions.html>
- pro vytvoření DC z MODS formátu musí být použito (a dle potřeb knihovny upraveno) oficiální mapování Kongresové knihovny – viz <http://www.loc.gov/standards/mods/mods-conversions.html>
- U digitalizovaných dokumentů je bibliografický popis vytvářen primárně z pohledu popisu fyzické předlohy, nejde o popis elektronického dokumentu

Zvukový dokument – fonografický váleček

- základní intelektuální entitou pro popis je zvukový dokument, tj. v jednom METS záznamu, který bude obsahovat metadata a strukturu jednoho a více válečků, budou MODS záznamy k této kolekci
- zvukovým dokumentem se rozumí vrcholná úroveň popisu fonografického válečku, ať již se týká válečku jen s jednou stranou nebo souboru mnoha válečků. Zvukový dokument odpovídá bibliografickému záznamu v katalogu. metadata budou popisovat entity³:
 - **zvukový dokument (soundcollection)** – zvukový nosič nebo soubor zvukových nosičů, na kterém byla zvuková nahrávka (tj. skladby nebo soubor skladeb) vydávány jako celek/ komplet. Může obsahovat jeden či více válečků v jednom obalu nebo ve více obalech, které jsou popsány jedním bibliografickým záznamem v katalogu a jejichž obsah spolu souvisí nebo na sebe navazuje. Zvukový dokument je nadřazená entita pro všechny následující.
 - **Skladba (soundrecording)** - popisuje zvlášť každou zvukovou nahrávku, tvořící ucelené dílo. • **příloha (supplement)** - přílohou se rozumí volně vložená entita k intelektuální entitě, např. volně vložený booklet, obsah, reklamní či edukativní příloha apod.

³ Toto pořadí nevyjadřuje logickou strukturu dokumentu, ta je popsána dále v DMF.

Rozlišujeme dva druhy příloh:

- 1) **obal (cover)** - popis pro fyzický obal kolekce. Skenuje se zátka, obal válečku, a pokud je to přínosné, tak i spodek obalu válečku. Pomocí elementu <relatedItem> se odkazuje ke kolekci.
- 2) **obraz válečku (imgdisc)** - popis pro sken válečku. Skenuje se okruží válečku a taky okruží a zátka vedle sebe.

Přílohy se pomocí elementu <relatedItem> odkazují k nejvyšší popisované úrovni, tj. zvukovému dokumentu. 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 (např. soundcollection, soundrecording, supplement). V případě opakování částí stejného typu se bude opakovat tomu odpovídající počet MODS záznamů v jednom SIP balíčku.
- každý MODS záznam má vlastní <dmdSec> část
- každý MODS záznam bude uložen ve vlastní METS části <dmdSec> pomocí mdWrap
- všechny top elementy MODS formátu jsou opakovatelné, kromě <recordInfo>
- všechny elementy Dublin Core jsou opakovatelné
- každá část <dmdSec> musí mít ID a vnořený element <mdWrap> s atributy MDTYPE a MIMETYPE.

Element	Atributy	Popis	Povinnost
<dmdSec>		Identifikátor dmdSec části METS záznamu	M
	ID	Pro <dmdSec> s popisem zvukového dokumentu hodnota "MODSMD_SOUNDCOLLECTION" a "DCMD_SOUNDCOLLECTION"	M
		Pro <dmdSec> s popisem skladby hodnota "MODSMD_SOUNDRECORDING" a "DCMD_SOUNDRECORDING"	M

		Pro <dmdSec> s popisem přílohy hodnota "MODSMD SUPPL" a "DCMD SUPPL"	M
<mdWrap>		Element obsahující vložené záznamy MODS	M
	MDTYPE	Hodnota „MODS“ pro záznamy v MODS, hodnota „DC“ pro záznamy v Dublin Core	M
	MDTYPEVERSION	Číslo verze MODS, hodnota pro záznamy v MODS, např. „3.6“	O
	MIMETYPE	Hodnota „text/xml“	O
<xmlData>			M

Pole MODS a Dublin Core pro jednotlivé úrovně zvukových dokumentů

- ID u elementu <mods>: Identifikátory budou začínat prefixy MODS_SOUNDCOLLECTION, MODS_SOUNDRECORDING, 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:

1) zvukový dokument :

- MODS_SOUNDCOLLECTION_0001
- DC_SOUNDCOLLECTION_0001

2) skladba:

- MODS_SOUNDRECORDING_0001
- DC_SOUNDRECORDING_0001

3) přílohy (skeny obalu a okruží):

- MODS_SUPPL_XXXX, kde XXXX je pořadové číslo přílohy, např. MODS_SUPPL_0001 je první příloha atd.
- DC_SUPPL_XXXX, kde XXXX je pořadové číslo přílohy, např. MODS_SUPPL_0001 je první příloha atd.

7.4.1. Pole MODS a DC pro úroveň zvukového dokumentu

Obsah sloupce "Popis":

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

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. pro úroveň zvukového dokumentu „MODS_SOUNDCOLLECTION_0001“	M	
<titleInfo>		Název kolekce, souborný název (pro plnění lze použít katalogizační záznam), samozřejmě lze využít všech prvků a elementů MODS, které názvové informace popisují) Opakovatelný element	M	
	type	hlavní název bez type-pole 245 \$a Může existovat pouze jeden hlavní název bez type. type: alternative-pole 246 translated -pole 242 uniform -pole 130, resp. 240 abbreviated – pole 210	MA	
<title>		Názvová informace	M	<dc:title>
<subTitle>		Podnázev	MA	<dc:title>

<partNumber>		Číslo části zvukového dokumentu, pokud je dělen na části. Údaj se čerpá z polí 240 a 245, podpole \$n.	MA	<dc:title>
<partName>		Název části z pole 240 nebo 245, podpole \$p.	MA	<dc:title>
<nonSort>		Část názvu, která má být vynechána při vyhledávání (např. <nonSort>The </nonSort>)	O	
<name>		Údaje o odpovědnosti. Přebíhá se z 1XX a 7XX marc21. Opakovatelný element pro více autorů/různé role	MA	
	type	Použit jednu z hodnot: -personal -corporate -conference -family	MA	
<namePart>		Údaje o křestním jméně a příjmení apod. Pokud je to možné, tak vyjádřit pro křestní jméno i příjmení. Více křestních jmen se zapíše do jednoho elementu <namePart>. Pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart>. Pokud známe datum narození a úmrtí autora, vyplnit ve tvaru RRRR-RRRR s atributem type="date".	MA	<dc:creator> Nutno do jednoho pole DC spojit jméno i příjmení
	type	Použit jednu z hodnot: -date (MA) -family (MA) -given (MA) -termsOfAddress (R) Pokud se nejedná o osobu, atribut type se nepoužije.		

<nameIdentifier>		Číslo národní autority	RA	
<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. Příklad: <mods:name> <etal>akol.</etal> </mods:name>	O	
<role>		Specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		Popis role – nutno použít kontrolovaný slovník, např. http://www.loc.gov/marc/relators/relaterm.html	M	
	type	Hodnota „code“	M	
	authority	Údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority=“marcrelator“	M	
<typeOfResource>		Hodnota „sound recording“, „sound recording–musical“ nebo „sound recording–non musical“ Mělo by se vyčítat z MARC21 katalogizačního záznamu, z pozice 06 návěští – hodnoty i, j	R	<dc:type>

<genre>		Bližší údaje a typu dokumentu Stanovená hodnota „sound recording“	M	<dc:type>model:soundrecording</dc:type>
<originInfo>		Informace u původu předlohy	M	
<place>		Údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	<dc:coverage>
<placeTerm>		Konkrétní určení místa vydání, např. Praha, resp. „xr“ pro ČR Odpovídá hodnotám z katalogizačního záznamu, pole 260 \$a, resp. pole 008/15-17	MA	<dc:coverage>
	type	Hodnota „code“ pro údaj z pole 008 Hodnota „text“ pro údaj z pole 260	M	
	authority	Hodnota „marccountry“ jen u údaje z pole 008	MA	
<publisher>		Jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala, odpovídá poli 260 \$b v MARC21 Pokud má titul více vydavatelů, přebírají se ze záznamu všichni (jsou v jednom poli 260)	MA	<dc:publisher>
<dateIssued>		Datum vydání předlohy, odpovídá hodnotě z katalogizačního záznamu, pole 260 \$c a pole 008/07-10	M	<dc:date>
	encoding	Hodnota „marc“ jen u údaje z pole 008	R	
	point	Hodnoty „start“ resp. „end“ jen u údaje pro rozmezí dat z pole 008	MA	
	qualifier	Možnosti dalšího upřesnění, hodnota „approximate“ pro data, kde nevíme přesný údaj	R	

<edition>		Údaje o pořadí vydání, odpovídá poli 250 \$a katalogizačního záznamu	R	
<issuance>		Údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07 Hodnota „monographic“	M	
<physicalDescription>		Obsahuje údaje o fyzickém popisu zdroje předlohy, pole 300, 306, 337, 338)	M	
<form>		Údaje o fyzické podobě dokumentu Hodnota „sound recording“ z údaje z 007/00 code s Hodnota „cylinder“ u údaje z 007/01 code e	M	<dc:format>
	authority	Hodnota „marccategory“ u údaje z 007/00 Hodnota „marcsmd“ u údaje z 007/01	MA	
<extent>		Údaje o rozsahu, odpovídá hodnotám v poli 300 \$a, \$b, \$c, \$e; 306 Příklad: 300 \$a 1 fonováleček 300 \$b analog, 160ot/m, mono 300 \$c průměr 54 mm, délka 106 mm	M	<dc:coverage>
<note>		Slouží k vepisování údajů z polí 5XX, 245 \$c. Pro každou poznámku musí být samostatný <note> element.	RA	<dc:description>
	type	Pro obsah pole 245 hodnota „statement of responsibility“ Pro obsah pole 508 hodnota „creation/production details“ Pro obsah pole 511 hodnota „performers“	MA	

		Pro obsah pole 518 hodnota „venue“ Pro obsah pole 546 hodnota „language“ Pro obsah pole 561 hodnota „ownership“ Pro obsah pole 562 hodnota „version identification“ Pro obsah pole 583 hodnota „action“		
<tableOfContents>		slouží k vepsání názvů skladeb na válečku; obsah 505\$a	MA	<dc:description>
	displayLabel	ind1=0 hodnota „Contents“ ind1=1 hodnota „Incomplete contents“ ind1=2 hodnota „Partial contents“	R	
<relatedItem>		Informace o dalších dokumentech, jejich částech či zdrojích, které jsou ve vztahu k popisovanému zvukovému dokumentu; Dále lze použít pro vyjádření edice, ve které je dokument vydán (tzn. kolekce je součástí nějaké edice), údaj o edici musí obsahovat minimálně element <title> s jejím názvem Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsány pro tyto elementy	RA	
	type	Např. hodnota „series“	MA	
<identifier>		Údaje o identifikátorech, obsahuje unikátní identifikátory nebo lokální	M	<dc:identifier>

	type	Budou se vyplňovat následující hodnoty, pokud existují: <ul style="list-style-type: none"> - issue number (MA) - uuid (M) - urn:nbn (M) - upc (MA) - local (MA) - music-publisher (MA) 	M	
<location>		Údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
<physicalLocation>		Údaje o instituci, kde je fyzicky uložen popisovaný dokument, např. MZK; nutno použít kontrolovaný slovník – sigly knihoven (BOA001); odpovídá poli 040 v MARC21 Pozn. U dokumentů ze soukromých sbírek není možné vyplnit	MA	<dc:source>
	authority	Hodnota „siglaADR“	M	
<shelfLocator>		Signatura nebo lokační údaje o dokumentu, který slouží jako předloha	M	<dc:source>
<recordInfo>		Údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		Popis standardu, ve kterém je přebíraný katalogizační záznam Odpovídá hodnotě návěští záznamu MARC21, pozice 18 – hodnota „aacr“, tj. pro LDR/18=“a“ bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA	MA	
<recordContentSource>		Kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	Hodnota „marcorg“	R	

<recordCreationDate>		Datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	Hodnota „iso8601“	M	
<recordChangeDate>		Datum změny záznamu, na úroveň minut podle normy ISO 8601	R	
	encoding	Hodnota „iso8601“	M	
<recordIdentifier>		Identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	Hodnota se přebírá z katalogu pole 003	R	
<recordOrigin>		Údaje o vzniku záznamu; hodnoty „machine generated“ nebo „human prepared“	R	
<languageOfCataloging>		Jazyk katalogového záznamu	R	
<languageTerm>		Přebírá se z katalogu – pole 040 \$b	R	
	authority	Hodnota „iso639-2b“	R	

7.4.2 Pole MODS a DC pro úroveň skladby

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	obsahuje název úrovně, pro úroveň skladby je to "MODS_SOUNDRECORDING_0001", kde číslo je pořadové číslo skladby v rámci zvukového dokumentu.	M	
<titleInfo>		názvové informace - pokud nemá skladba zjištělný název, kopíruje se informace z nadřazené struktury	M	

	type	alternative uniform translated	MA	
<title>		hlavní název	M	<dc:title>
<subTitle>		podnázev	MA	<dc:title>
<partNumber >		číslo skladby či její pořadí	MA	<dc:title>
<partName>		Název části z pole 240 nebo 245, podpole \$p.	MA	<dc:title>
<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 corporate conference family	MA	

<namePart>		<p>Údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení. Více křestních jmen se zapíše do jednoho elementu <namePart>. Pokud nelze rozlišit křestní jméno a příjmení, nepoužije se type 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 type="date".</p>	MA	<dc:creator>
	type	<p>Použít jednu z hodnot:</p> <p>date (MA)</p> <p>family (MA)</p> <p>given (MA)</p> <p>termsOfAddress (R)</p> <p>Pokud se nejedná o osobu, atribut type se nepoužije.</p>	MA	
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
	type	<p>hodnota "code" – kód role z kontrolovaného slovníku (http://www.loc.gov/marc/relators/relaterm.html)</p>	M	
	authority	<p>údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"</p>	M	

<genre>	type	stanovená hodnota „sound recording“	M	<dc:type>model:soundrecording</dc:type>
<originInfo>		Informace o původu předlohy	M	
<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	<dc:coverage>
<placeTerm>		konkrétní určení místa vydání, např. Praha, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 260 \$a, pole 044 \$c, resp. pole 008/15-17	MA	<dc:coverage>
	type	hodnota "code" pro údaj z pole 008 hodnota "text" pro údaj z pole 260	M	
	authority	hodnota "marccountry" jen u údaje z pole 008	MA	
<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07 hodnota "monographic"	M	

<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy (pole 300, 306)	M	
-----------------------	--	--	---	--

<extent>		údaje o rozsahu; odpovídá hodnotám v poli 300, 306 \$a	M	<dc:coverage >
<language>		údaje o jazyce dokumentu v případě vícenásobného výskytu nutno element <language> opakovat	M	
	objectPart	možnost vyjádřit jazyk konkrétní části díla; možné hodnoty: - sung or spoken text, odpovídá poli 041 \$d - translation, odpovídá poli 041 \$h - libretto, odpovídá poli 041 \$e -accompanying materiál, odpovídá poli 041 \$g	M A	
<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 >
	type	hodnota „code“	M	
	authority	hodnota „iso 639-2b“	M	
<typeOfResource>		Obsahuje hodnotu z návěstí z pozice 05: - „sound recording-musical“ – odpovídá \$j - „sound recording-nonmusical“ – odpovídá \$i	R	

<note>		<p>Obecná poznámka, může též sloužit jako element pro vyjádření údaje o odpovědnosti, poznámka o realizátorech, účinkujících či jazyku.</p> <p>Pro každou poznámku je potřeba samostatný element <note></p>	R A	<dc:description>
	type	<ul style="list-style-type: none"> - statement of responsibility“ – odpovídá poli 245 \$c - „creation/ production credits“ – odpovídá poli 508 \$a - „performers“ – odpovídá poli 511 \$a - „venue“ – odpovídá poli 518 \$a - „language“ – odpovídá poli 546 \$a - version identification – odpovídá poli 562 \$b - action – odpovídá poli 583 \$a 	M	
<subject>		<p>údaje o věcném třídění, předpokládá se přebírání z katalogizačního záznamu</p>	R	
	authority	<p>„vyplnit hodnotu „czenas“, „eczenas“ z pole 072 \$2, nebo jinou hodnotu, z tohoto pole. Případně lze ponechat prázdné“.</p>	M	

<topic>		libovolný výraz specifikující nebo charakterizující obsah titulu; použit 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>
<geographic>		geografické věcné třídění použit 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>
<temporal>		chronologické věcné třídění použit 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>
<name>		jméno použité jako věcné záhlaví použit 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; struktura a atributy stejné jako pro údaje o původcích – viz element <name>	R	<dc:subject>
<title>		Pole použité jako vedlejší věcné záhlaví	R	<dc:subject>

<titleInfo>		Unifikovaný název	M	
<genre>		Použit' při 653 a ind2=6	R	<dc:subject>
<geographicCode>		Kód geografické oblasti, obsah pole 043	R	<dc:subject>
	authority	Hodnota „marc9ac“	M	
<relatedItem>		<p>Informace o zbukovém dokumentu, jehož je skladba součástí</p> <p>Údaje o zvukovém dokumentu musí obsahovat nejméně element <title> s jeho názvem.</p> <p>Poznámka: element <relatedItem>. Může obsahovat jakýkoliv jiný elementn MODS – jejich použití se řídí pravidly popsanými pro tyto elementy.</p>	R A	
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální	M	<dc:identifier>
	type	<p>budou se povinně vyplňovat následující hodnoty, pokud existují:</p> <ul style="list-style-type: none"> - uuid (M) - „issue number“ – číslo válečku, v MARC21 pole 028 ind1=0 \$a, např. \$a DV 6196, přebírá se z katalogu, ind1=0 (MA) - type=“matrix-number“ pro unikátní matriční číslo konkrétní 	M	

		<p>nahrávky, pole 028 ind1=1 \$a (MA)</p> <p>- music-publisher 028a, b ind1=3 (R)</p> <p>- upc, ean (R)</p> <p>- jiný interní identifikátor - hodnota atributu „local“ (R)</p>		
<recordInfo>		<p>údaje o metadatovém záznamu válečku – jeho vzniku, změnách apod.</p>	M	
<descriptionStandard>		<p>popis standardu, ve kterém je přebíraný katalogizační záznam odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota “acr”, tj. pro LDR/18=”a” - bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA</p>	M A	
<recordContentSource>		<p>kód nebo jméno instituce, která záznam vytvořila nebo změnila</p>	R	
	authority	<p>Hodnota „marcorg“</p>	R	

<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"	M	
<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 na úroveň minut	R	
	encoding	hodnota atributu "iso8601"	M	
<recordOrigin>		údaje o vzniku záznamu; hodnoty: "machine generated" nebo "human prepared"	R	
<languageOfCataloging>		jazyk katalogového záznamu	R	
<languageTerm>	authority	přebírá se z katalogu - pole 40 \$b Hodnota „iso639-2b“	R R	

7.4.3 Pole MODS a DC pro obrazové informace (přílohy)

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	
<titleInfo>		názvové informace o příloze	M	
<title>		uazapíše se, která část fonoválčku je na skenu; hodnoty: „Zátka“, „Obal“, „Spodek obalu“, „Okružní“, „Zátka a okružní“	MA	<dc:title>
<typeOfResource>		Možné hodnoty „still image“	M	
<genre>		Hodnota „supplement“	M	<dc:type>model:supplement</dc:type>
	type	Na určení typu přílohy Povolené hodnoty: cover, imgdisc	M	

<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy	M	
<form>		Údaje o fyzické podobě dokumentu Hodnoty se přejímají z kontrolovaných slovníků		
	authority	Hodnota „marccategory“	MA	<dc:format>
<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	
<note>		obecná poznámka k příloze	RA	<dc:description>
<relatedItem>		Popis zvukového dokumentu, ke kterému příloha patří. Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsánymi pro tyto elementy	M	
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální – viz přehled typů atributů níže	M	<dc:identifier>

	type	budou se povinně vyplňovat následující hodnoty, pokud existují: - uuid (M)	M	
<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "aacr", tj. pro LDR/18="a" - bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA	MA	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila; nutno vytvořit kontrolovaný slovník	R	
<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"	M	
<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 na úroveň minut	R	
	encoding	hodnota atributu "iso8601"	M	

<recordOrigin >		údaje o vzniku záznamu; hodnoty „machine generated“ nebo „human prepared“	R	
<languageOfC ataloging>		jazyk katalogového záznamu	R	
<langua geTerm>		přebírá se z katalogu - pole 40 \$b	R	
	auth ority	hodnota “iso639-2b”	R	

7. 5 Technická a administrativní metadata ve vedlejších METS souborech

- technická a administrativní metadata přicházející z digitalizace jsou dále v maximální míře ukládána v LTP systému (po konverzi do interního formátu LTP systému)
- Technická a administrativní metadata se zapisují do vedlejších souborů METS (amd_mets.xml)
- tento vedlejší METS záznam (amd_mets.xml) je linkován z hlavního METS záznamu dokumentu (v části <fileSec>)
- Technická a administrativní metadata vznikají pro každý soubor archivních zvukových dat (pro SA i MCA) a pro každý soubor archivních obrazových dat
- technická a administrativní metadata u obrazových souborů budou vznikat i pro prvotní sken (většinou TIFF), který se po nutných úpravách maže a dále neuchovává
- Pro popis zvukových dat se použijí metadatové formáty PREMIS (část OBJECT) a AES57
- Pro popis obrazových dat se použijí metadatové formáty PREMIS (část OBJECT) a MIX
- Pro popis událostí a souvisejících agentů v procesu vzniku digitálních kopií se používají části metadatového schématu PREMIS—PREMIS EVENT a PREMIS AGENT, pro zvuková data je vloženo lokální schéma ndktech do části PREMIS AGENT. Zachycují veškeré změny a procesy, které byly na datech i metadatach provedeny
- Každý archivovaný zvukový soubor bude tedy ve vedlejším METS záznamu obsahovat tyto části a metadatová schémata: PREMIS Object, AES57, PREMIS Event, PREMIS Agent vše zabalené do kontejnerového schématu METS. Každý archivovaný obrazový soubor bude ve vedlejším METS záznamu obsahovat tyto části a metadatová schémata: PREMIS Object, MIX, PREMIS Event, PREMIS Agent vše zabalené do kontejnerového schématu METS.
- technická a administrativní metadata budou zabalena v části <amdSec> formátu METS ve vlastních formátech (AES57, MIX, PREMIS – části object; events; agent)
- <amdSec> část bude existovat vždy jedna pro všechny reprezentace jedné zvukové stopy či obrazové stránky dokumentu (MC, UC, u obrazu ALTO XML a OCR.txt) a bude obsahovat metadata v <techMD> a <digiprovMD> podčástech
- část <amdSec> musí mít ID a vnořený element <techMD> nebo <digiprovMD>, oba s ID a vnořeným elementem <mdWrap> s atributem MDTYPE
- technická metadata jsou určena primárně pro zachycení technických informací o formátech souborů, o výsledcích validací a kontrol
- plnění technických metadat se předpokládá z výstupů vzniklých využitím služeb třetích stran jakou jsou JHOVE2, PRONOM aj.

Element	Atributy	Popis	Povinnost
<amdSec>		element obsahující technická metadata ve formátu PREMIS, AES57 nebo MIX	M

	ID	ID – identifikátor konkrétní části <amdSec>, např. pro jeden zvukový soubor „SOUND001“, obrazovou stránku 1 hodnota “PAGE0001”, pro obrazovou stránku 2 hodnota “PAGE0002” atd.	
<techMD> nebo <digiprovMD>		element rozlišující typy jednotlivých administrativních metadat	M
	ID	ID pro část <techMD>:	
		<p>pro části obsahující PREMIS Object hodnota:</p> <ul style="list-style-type: none"> • “OBJ_001” pro původní smazaný soubor TIFF u obrazu nebo WAV (SA) u zvuku • “OBJ_002” pro MC u obrazu a MCA u zvuku • “OBJ_003” pro ALTO XML u obrazu <p>počet PREMIS Object není omezen, číslování pokračuje OBJ_004 atd.</p> <p>pro části obsahující AES57 hodnota:</p> <ul style="list-style-type: none"> • “AES57_001” pro původní zvukový soubor • “AES57_002” pro MCA • “AES57_003” pro druhou MCA <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>	
		ID pro část <digiprovMD>:	
		<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. 	

<mdWrap>		element obsahující vložené záznamy PREMIS, AES57 či MIX	M
	MDTYPE	pro záznamy PREMIS object, event i agent vždy hodnota "PREMIS" pro záznamy AES57 hodnota "AES57" pro záznamy MIX hodnota "NISOIMG"	M
<xmlData>			M

7.5.1 PREMIS Object

- popisovat se pomocí PREMIS Object budou soubory, tj. dle specifikace PREMIS vždy úroveň tzv. **file** (ne reprezentace ani bitstream)
- záznam v PREMIS Object se bude vytvářet pro každý soubor:
 - vzniklý v procesu digitalizace - původní zvukový soubor, který se archivuje, a původní obraz, který se maže
 - archivní kopie zvukové stopy
 - archivní kopie obrazové stránky
 - ALTO XML u obrazu,
 - případné další reprezentace zvukové stopy / obrazové stránky (například nová archivní kopie vytvořená migrací z původní MC)
- PREMIS Object se nebude vytvářet pro UC (zvuk i 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 archivní soubor WAV je tímto způsobem nalinkován na původní nahrávku (resp. na jeho PREMIS Object záznam) – pomocí tagu <relatedObjectIdentification>, který obsahuje ID původního objektu
 - zároveň pomocí tagu <relatedEventIdentification> je záznam PREMIS Object archivního souboru WAV 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) a XML (ALTO)

Element	Popis	Použití pro	Povinnost
<object>	kořenový element pro PREMIS Object; použít vždy s atributem dle typu objektu xsi:type="file" – pro soubor 1-1	vše	M
<objectIdentifier>	identifikátor k jednoznačnému odlišení objektu v určitém kontextu; 1-n	MC, XML, PS	M
<objectIdentifierType>	popis kontextu, ve kterém je identifikátor unikátní, např. NDK, ANL nebo název repozitáře; nutno použít kontrolovaný slovník; 1-1	MC, XML, PS	M
<objectIdentifierValue>	vlastní hodnota identifikátoru, např. img0001-master, urn.nbn.cz-123465 apod.; 1-1	MC, XML, PS	M

<preservationLevel>	údaje o úrovni ochrany souboru, která se na něj vztahuje; některé soubory nejsou tak důležité jako jiné, mají menší úroveň ochrany; 0-n	MC, XML, PS	MA
<preservationLevelValue>	hodnota úrovně ochrany, která je pro soubor relevantní, pro původní sken PS hodnota deleted, pro MC a XML hodnota preservation; 1-1	MC, XML, PS	M
<preservationLevelDateAssigned>	datum, kdy byla přiřazena hodnota úrovně ochrany, zápis v ISO 8601, na úroveň dne (RRRR-MM-DD) 0-1	MC, XML, PS	R
<objectCharacteristics>	technické údaje o souboru 1-n	MC, XML, PS	M
<compositionLevel>	údaj o tom, zda je nutné digitální objekt rozbalit nebo dekodovat; např. 0 (defaultně pro žádné zabalení nebo kódování); 1 pro jedno zabalení a kódování, podobně pak hodnota 2; 1-1	MC, XML, PS	M
<fixity>	údaje o kontrolním součtu 1-n	MC, XML, PS	M
<messageDigestAlgorithm>	použitý algoritmus kontrolního součtu, např. MD5 aj. 1-1	MC, XML, PS	M
<messageDigest>	hodnota kontrolního součtu 1-1	MC, XML, PS	M

<messageDigestOriginator>	agent (osoba, instituce, stroj, SW), který kontrolní součet vytvořil (např. JHOVE apod.) 0-1	MC, XML, PS	M
<size>	údaje o velikosti souboru v bytech 1-1	MC, XML, PS	M
<format>	údaje o formátu souboru pro soubory ALTO XML je možné vytvořit element dvakrát, jednou popisuje formát XML, podruhé obsahuje informace o použitém standardu ALTO 2.0 (viz příklad 2) 1-n	MC, XML, PS	M
<formatDesignation>	identifikace formátu souboru, výstup z JHOVE, PRONOM služeb apod. 1-1	MC, XML, PS	M
<formatName>	jméno formátu dle MIME type, audio/wav 1-1	MC, XML, PS	M
<formatVersion>	verze formátu, např. 6.0 nebo jméno verze (př. PCMWAVEFORMAT--vypíše např. JHOVE) 1-1	MC, XML, PS	MA
<formatRegistry>	identifikace formátu – dodatečná informace o záznamu formátů v registrech formátů (PRONOM) 1-1	MC, XML, PS	M
<formatRegistryName>	jméno použitého registru formátů, vždy PRONOM 1-1	MC, XML, PS	M

<formatRegistryKey>	unikátní identifikátor (označení) formátu v registru, povinné použít PUID z registru PRONOM, další registry možné 1-n	MC, XML, PS	M
<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 JPEG2000, sw pro úpravu zvuku každá aplikace zvlášť v novém elementu 1-n	MC, XML, PS	M
<creatingApplicationName>	název aplikace, např. WaveLab, ImageGear, Kakadu apod.;; 1-1	MC, XML, PS	M
<creatingApplicationVersion>	verze aplikace, např. 15.03.000 1-1	MC, XML, PS	M
<dateCreatedByApplication>	datum a čas vytvoření, např. 2008-11-10T12:37:46; musí být ve tvaru ISO 8601 (na úrovni vteřin); např. 2008-11-10T12:37:46; 1-1	MC, XML, PS	M
<originalName>	původní jméno souboru, např. 123456789_01.wav 1-1	MC, XML, PS	M
<relationship>	vyjádření vztahu popisovaného souboru k jiným souborům a událostem (eventům) 1-n	MC, XML	M

<relationshipType>	typ vztahu, doporučené hodnoty: derivation = vztah, kde objekt je výsledkem změny jiného objektu; structural = vztah mezi částmi objektu; tj. např. MC vytvořené z SA bude mít vztah “derivation”; 1-1	MC, XML;	M
<relationshipSubType>	upřesnění vztahu, doporučené hodnoty: created from; has source; is source of; has sibling; has part; is part of; has root; includes; is included in apod.; tj. např. MC vytvořené z SA budou mít vztah “created from”; 1-1	MC, XML;	M
<relatedObjectIdentification>	identifikace souvisejícího souboru 1-n pro MC, XML pro vyjádření vztahu k původnímu objektu	MC, XML	M
<relatedObjectIdentifierType>	specifikace kontextu, ve kterém je identifikátor souboru jedinečný, např. URN; temporary filepath; objectID 1-1	MC, XML	M
<relatedObjectIdentifierValue>	vlastní řetězec identifikátoru, např. URN:NBN:cz-1301091_011#0001 nebo název souboru, cesta k souboru apod. 1-1	MC, XML	M
<relatedEventIdentification>	identifikace s popisovaným souborem související události (eventu); seznam událostí viz PREMIS event 1-n	MC, XML	M
<relatedEventIdentifierType>	typ události, např. interní číslovací systém událostí jako no.nb.evt; NK repository event ID, UUID apod.	MC, XML	M

	1-1		
<relatedEventIdentifierValue>	hodnota identifikátoru události, např. NK_EVT_005 nebo hodnota UUID aj. 1-1	MC, XML	M
<relatedEventSequence>	pořadí události, např. 003; k určení pořadí lze určit datum události 0-1	MC, XML	R
<linkingEventIdentifier>	identifikátor události týkající původního skenu PS; typy událostí mohou být např. vytvoření, smazání 1-n pro PS nutný link na události vytvoření (digitalizace) a jeho vymazání	PS	M
<linkingEventIdentifierType>	typ identifikátoru události, např. UUID, NK_eventID, vlastní číslovací systém apod. 1-1	PS	M
<linkingEventIdentifierValue>	hodnota identifikátoru, např. event_01; img0001-master-event001 apod. 1-1	PS	M

Archivační, ochranná metadata

7.5.2 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 zvukových a obrazových datech:
 - digitalizace – vytvoření prvního skenu / nahrávky SA
 - vygenerování MC obrazu, případně i zvukového souboru
 - 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) v jeho části <amdSec>, podčást <digiprovMD>
 - amd_mets.xml je linkován z hlavního METS záznamu dokumentu (v částí <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

Pole záznamu PREMIS Event

Element	Popis	Povinnost
<eventIdentifier>	údaje o identifikátoru události v kontextu digitalizace nebo repozitáře 1-1	M
<eventIdentifierType>	typ identifikátoru, např. no.nb.evt; NK_eventID, UUID apod. 1-1	M
<eventIdentifierValue>	hodnota identifikátoru, např. EVT_001; event_019 apod.	M

	1-1	
<eventType>	kategorizace události, nutno použít kontrolovaný slovník; typy událostí, které musí být zaznamenány: capture, migration, deletion 1-1	M
<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 /: <ul style="list-style-type: none"> - capture/digitization – vznik prvního skenu/nahrávky - capture/XML_creation - capture/TXT_creation - migration/MC_creation - deletion/PS_deletion 1-1	M
<eventOutcomeInformation>	informace o výsledku události 0-n	R
<eventOutcome>	kategorizace výsledku události, např. slovy jako successful, failure; možno použít kódy dle kontrolovaného slovníku nebo seznamu kódů; 1-n	M
<linkingAgentIdentifier>	identifikace jednoho nebo více agentů spojených s událostí 1-n	M
<linkingAgentIdentifierType>	označení typu identifikátoru, např. NK_AgentID, UUID apod. 1-1	M

<linkingAgentIdentifierValue>	hodnota identifikátoru, např. agent_softwareName_5.2; agent_novakJ apod. 1-1	M
<linkingAgentRole>	role agenta ve vztahu k události, např. software; SW component; operator; 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
<linkingObjectIdentifier>	informace o objektu/souboru spojeného s událostí, link na něj 0-n	M
<linkingObjectIdentifierType>	označení typu identifikátoru, např. hodnota "file" 1-1	M
<linkingObjectIdentifierValue>	hodnota identifikátoru, např. URN:NBN:cz-_0011#0001 aj. nebo PS_1_0003_R 1-1	M

7.5.3 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
 - pro fonoválčky se bude zaznamenávat následující zařízení: cylinder player, předzesilovač, ADA konvertor, editační nástroj pro zvuk, hrot, tone arm
- 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) v jeho části <amdSec> a podčásti <digiprovMD>
 - amd_mets.xml je linkován z hlavního METS záznamu dokumentu
- pro každého agenta, tj. jeden PREMIS Agent záznam, bude vytvořena jedna <digiprovMD> část

Obsah pole „Popis“:

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

Element	Popis	Povinnost
<agentIdentifier>	popis identifikátoru, který jednoznačně označuje agenta v rámci jednoho kontextu (repozitář např.) 1-n	M
<agentIdentifierType>	označení typu identifikátoru, např. NK_AgentID, UUID apod. 1-1	M
<agentIdentifierValue>	hodnota identifikátoru, např. agent_softwareName_5.2; agent_novakJ apod. 1-1	M
<agentName>	textové upřesnění agenta, např. přesný název SW, plné jméno osoby apod. - FixImage1.3; Jan Novák; CCS docWorks 6.2.1;	R

	0-n	
<agentType>	obecné označení agenta – pro osoby např. osoba, pro SW např. software apod. hodnoty: organization; person; software, hardware 1-1	M
<agentNote>	použití pouze pokud je <agentType> Software a půjde o agenta souvisejícího s migrací původního souboru např. TIFF na JPEG2000 (creation/migration Event); bude obsahovat příkaz k výrobě JPEG2000 souboru v aplikaci (např. aplikace Kakadu, OpenJpeg) 1-n	MA
<agentExtension>	element je určen pro popis použitého zařízení u zvukových dokumentů (zvl. hardware)	MA
<NK:manufacturer>	výrobce (př. Mytek, Sony)	MA
<NK:serialNumber>	sériové číslo zařízení (př. 01504-1208-043)	M
<NK:settings>	nastavení zařízení, volný text	R
<NK:functionalRole>	Zapíše se role agenta v procesu digitalizace nosiče. Výběr z hodnot: Reproducer, Amplifier, ADC, Recorder, Digital audio editor	R

7.5.4 Technická metadata pro zvukové soubory

Pro zápis technických metadat zvukových souborů se používá metadatové schéma AES57. Povinné je popsat digitální soubor (dle povinností uvedených níže). Ve vlastní metadatové části se popisuje se i původní analogový zdroj, tj. fonováleček.

7.5.4.1 Pro digitální dokument

element	atributy	popis	Povinnost
<audioObject>			M
	<i>ID</i>		M
	<i>title</i>	Např. název dokumentu	R
	<i>analogDigitalFlag</i>	pro digitální dokument vždy hodnota FILE_DIGITAL	M
	<i>schemaVersion</i>	Verze standardu AES57, vždy hodnota "1.0.0"	M
	<i>disposition</i>	Informace o aktuálním stavu objektu, resp. o jeho uložení, umístění, například „Deposited into DMS“, „Unknown“	M
<format>		Oficiální název formátu, tj. „Waveform_Audio“ (nastavit fixně pro soubory WAV) 1-1	M
	<i>specificationVersion</i>	verze formátu, např. „1 PCM encoding“	MA

<appSpecificData>		Aplikace, která se podílí na tvorbě audia a která do souboru vložila nějaká svá metadata 0-n	O
	<i>appVersion</i>	Verze aplikace	M
<audioDataEncoding>		Použitá modulační metoda převodu analogového signálu na signál digitální. (Pro WAV soubory dle NDK hodnota PCM, příp. PCM audio in integer format apod.) 1-1	M
<byteOrder>		Endianita, pořadí bajtů, možné hodnoty: LITTLE_ENDIAN nebo BIG_ENDIAN 1-1	M
<firstSampleOffset>		Počet bajtů, které se v souboru vyskytují ještě před popisovaným audio objektem (před prvním bajtem audio dat). Hodnotu vypíše charakterizační nástroje, např. FITS a JHOVE. 0-1	RA
<audioDataBlockSize>		Velikost datových bloků popisovaného audio dokumentu, velikost se udává v bajtech. Zapíše se číslo, např. „3“. 0-1	RA

<use>		Účel popisovaného souboru. 1-n	M
	<i>useType</i>	Povolené hodnoty: PRESERVATION_MASTER	M
<primaryIdentifier>		Údaj o identifikátoru popisovaného objektu. 1-1	M
	<i>identifierType</i>	Může být jen: UMID nebo FILE_NAME nebo SHELF_NUMBER nebo OTHER	M
	<i>idOtherType</i>	Vyplníme, pokud u atributu <i>identifierType</i> zapíšeme OTHER, zde může být cokoliv	RA
<secondaryIdentifier>		Údaj o dalším identifikátoru popisovaného objektu. Element může být použit opakovaně.	O
	<i>identifierType</i>	Může být jen: UMID nebo FILE_NAME nebo SHELF_NUMBER nebo OTHER	M
	<i>idOtherType</i>	Vyplníme, pokud u atributu <i>identifierType</i> zapíšeme OTHER, zde může být cokoliv, například ACCESSION NUMBER apod.	RA
<face>		Úroveň audioObject má nejméně jednu úroveň face. Face obsahuje jeden nebo více kanálů (streams), úroveň face zpravidla odpovídá jednomu WAV souboru. Nástroje JHOVE a FITS rovnou vypíší jednotlivé úrovně (face, region, stream) charakterizovaného souboru.	M

		1-n	
	<i>ID</i>	Identifikátor této části dat, který se bude používat při odkazování z jiných částí záznamů (z region) Př. Side_1	M
	<i>direction</i>	u digitálního souboru vždy hodnota "NONE"	M
	<i>audioObjectRef</i>	Odkaz na ID audioObject	M
	<i>label</i>	Název logického audio celku, např. název strany desky, pokud nějaký je (Rusalka, Půlnoční....) nebo Side_A. Tento název se pak může zobrazit uživateli v sw aplikaci.	M
<timeline>		Element obsahující subelementy popisující délku nahrávky a čas, kdy začíná 1-1	M
<startTime>		Začátek nahrávky, např. 0, udává se ve frames 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného audio souboru, v herzích, např. "96000"	
<duration>		délka nahrávky, udává se ve frames Př. 55642500 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného audio souboru, v herzích, např. "96000"	M

<region>		Další úroveň popisovaného audio objektu. Jednotlivé části nahrávky (face), úroveň face má nejméně jednu úroveň region. Úroveň region může odpovídat jednotlivým stopám. 1-n	M
	<i>ID</i>	ID této úrovně se odvozuje od ID vyšší úrovně face. K ID úrovně face se přidá 0001, 0002 apod., dle počtu regionů.	M
	<i>formatRef</i>	Obsahem atributu je odkaz na atribut ID elementu formatRegion	M
	<i>faceRef</i>	Obsahem atributu je odkaz na ID atribut vyšší úrovně face	M
	<i>label</i>	Název této úrovně audio objektu, např. Region_1 nebo název písně, proslovu apod.	M
<timeRange>		Element obsahující subelementy popisující délku popisovaného regionu a čas, kdy začíná 1-1	M
<startTime>		Začátek regionu, např. 0, udává se ve frames 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<duration>		Délka regionu ve frames.	M

		1-1	
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. “96000”	M
<numChannels>		Počet kanálů (audio streams) regionu Zapísovanou hodnotou je číslo, př. „1“ 1-1	M
<conditionNote>		Tento element může být použit k popisu stavu popisovaného objektu. 0-n	O
<note>		text poznámky 1-1	M
<noteTimeRange>		Element se použije k zápisu času (začátku a trvání) pro výše uvedenou poznámku ke stavu objektu 0-1	O
<startTime>		Začátek části nahrávky, pro který je platná poznámka, např. 0, udává se ve frames 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. “96000”	M
<duration>		Délka po kterou je poznámky platná ve frames. 1-1	M

	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<creationDate>		Čas vytvoření poznámky, zápis dle normy ISO 8601 1-1	M
<securityNote>		Obsahem elementu je informace (volný text) o obsahu citlivých údajů (telefonních čísel, jmen...) 0-n	O
<stream>		Jednotlivé kanály, tj. u stereo nahrávky budou 2 elementy stream u mono nahrávky jen jeden element stream 1-n	M
	<i>ID</i>	Odvozuje se od ID vyšší úrovně (regionu), k id regionu se přidá písmeno, např. u008_20162011_021_0001a	M
	<i>label</i>	Pojmenování kanálu. např. „1“ nebo „Stream1“	M
	<i>faceRegionRef</i>	Obsahem atributu je odkaz na ID regionu.	M
<channelAssignment>		Obsahem elementu je popis pozice kanálu. 1-1	M
	<i>channelNum</i>	číselné označení popisovaného kanálu, př. „1“	M

	<i>leftRightPosition</i>	Číselné označení umístění kanálu na ose levá-pravá. Hodnoty např. „0.0“ (střed) nebo „100.0“ (vpravo) nebo „-100.0“ (vlevo)	O
	<i>frontRearPosition</i>	Číselné označení umístění kanálu na ose...	O
<conditionNote>		Poznámka o stavu popisovaného kanálu (stream) 0-n	O
<note>		Text poznámky 1-1	M
<noteTimeRange>		Časový rozsah části kanálu, které se poznámka týká 0-1	O
<startTime>		Začátek části nahrávky, pro který je platná poznámka, např. 0, udává se ve frames 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. “96000”	M
<duration>		Délka po kterou je poznámky platná ve frames. 1-1	M

	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<creationDate>		Čas vytvoření poznámky, zápis dle normy ISO 8601 1-1	M
<formatList>		Obsahuje soupis všech regionů v popisovaném audioObject 1-1	M
<formatRegion>		Obsahem elementu jsou souhrné a dodatečné informace o úrovni region. 1-n	M
	<i>ID</i>	Identifikátor části	M
	<i>label</i>	Použije se pro zápis mimetypeu, tj. „audio/x-wav“	M
	<i>ownerRef</i>	Soupis všech ID regionů	M
<bitDepth>		Bitová hloubka digitálního zvukového dokumentu. Viz Standard pro zvuková data na webu ndk.cz	M
<sampleRate>		Vzorkovací frekvence, např. „96000“. Viz Standard pro zvuková data na webu ndk.cz 1-1	M
<wordSize>		Velikost přenášeného vzorku v bajtech. Zapisovanou hodnotou je číslo, např. „3“ (odpovídá 24 bitové hloubce)	R

<soundField>		Obsahem elementu je popis zvukového pole, možné hodnoty jsou: „MONO“ nebo „STEREO“ nebo „SURROUND“	R

7.5.4.2 Pro původní zdrojový (analogový) dokument (př. gramodeska), bude v METS v části <sourceMD>

Ve schématu METS, v části sourceMD je možné (dobrovolné) popsat původní objekt (kniha, videokazeta, gramofonová deska). V případě fonografických válečků je možné tuto část metadat z velké části plnit namapováním na katalogizační záznam na pole 007 a 344 (RDA), tato pole jsou ovšem v katalogizačním záznamu nepovinná.

element	atributy	popis	Povinnost
<audioObject>		Pro popis původního analogového dokumentu	M
	<i>ID</i>		M
	<i>title</i>	název dokumentu	O
	<i>analogDigitalFlag</i>	Povolené hodnoty: ANALOG	M
	<i>generation</i>	Kolikátá je generace fyzického nosiče, u původního analogového obvykle hodnota "Original"	O
	<i>disposition</i>	kde je dokument uložen nyní, př. <i>returned to stacks, returned to National museum..</i>	M
	<i>schemaVersion</i>	Verze standardu AES57, vždy hodnota "1.0.0"	M

<format>		Obsahem elementu je název nosiče z kontrolovaného slovníku.	M
	<i>specificationVersion</i>	Verze nosiče, pokud existuje	RA
<physicalProperties>		Obsahem elementu je popis fyzických vlastností nosiče.	O
<cylinderStructure>		Element sloužící k popisu fyzických vlastností nosiče. Součástí jsou elementy pro popis jednotlivých vrstev nosiče, rozměrů apod.	O
	<i>stockBrand</i>	Výrobce nosiče, jméno produktu, např. „Supraphon“, „HMV“, „ESTA“	O
<substrateMaterialLayer>		Popis základního materiálu nosiče.	M
	<i>composition</i>	Složení vrstvy	O
	<i>role</i>	Role vrstvy, Může být: LABEL_LAYER, PROTECTIVE_LAYER,DATA_LAYER, PROTECTIVE_DATA_LAYER,SUPPORT_LAYER	M
	<i>order</i>	Pořadí vrstvy	M
<thickness>		Tloušťka této vrstvy	O
	<i>unit</i>	Jednotka tloušťky, např. „MICROMETRES“, „MILIMETRES“	M
<fillerLayer>		Popis vnitřní vrstvy nosiče.	O
	<i>composition</i>	Složení vrstvy	O
	<i>role</i>	Role vrstvy, Může být: LABEL_LAYER, PROTECTIVE_LAYER,DATA_LAYER,	M

		PROTECTIVE_DATA_LAYER,SUPPORT_LAYER	
	<i>order</i>	Pořadí vrstvy	M
<thickness>		Tloušťka této vrstvy	O
	<i>unit</i>	Jednotka tloušťky, např. „MICROMETRES“, „MILIMETRES“	M
<surfaceLayer>		Popis svrchní vrstvy nosiče	O
	<i>composition</i>	Složení vrstvy	O
	<i>role</i>	Role vrstvy, Může být: LABEL_LAYER, PROTECTIVE_LAYER,DATA_LAYER, PROTECTIVE_DATA_LAYER,SUPPORT_LAYER	M
	<i>order</i>	Pořadí vrstvy	M
<thickness>		Tloušťka této vrstvy	O
	<i>unit</i>	Jednotka tloušťky, např. „MICROMETRES“, „MILIMETRES“	M
<dimensions>		Obsahem elementu je popis rozměrů nosiče	M
<diameter >		Průmět válečku	O
	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<length>		Délka válečku	O

	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<shellDimensions>		Obsahem tohoto elementu je popis obalu audionosiče	O
<length>		Délka obalu	O
	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<width>		Šířka obalu	O
	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<depth>		Hloubka obalu	O
	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<diameter>		Průměr obalu	O
	<i>unit</i>	Jednotka rozměru, např. „MICROMETRES“, „MILIMETRES“, „CENTIMETRES“, „INCHES“ apod.	M
<use>		Účel souboru	M
	<i>useType</i>	Povolené hodnoty: ORIGINAL_MASTER, OTHER	
<primaryIdentifier>		Údaj o identifikátoru popisovaného analogového objektu, gramfonové desky	M

	<i>identifierType</i>	Může být jen: UMID nebo FILE_NAME nebo SHELF_NUMBER nebo OTHER	M
	<i>odOtherType</i>	Vyplníme, pokud u atributu <i>identifierType</i> zapíšeme OTHER, zde může být cokoliv	O
<secondaryIdentifier>		Údaj o dalším identifikátoru popisovaného objektu. Element může být použit opakovaně.	O
	<i>identifierType</i>	Může být jen: UMID nebo FILE_NAME nebo SHELF_NUMBER nebo OTHER	M
	<i>idOtherType</i>	Vyplníme, pokud u atributu <i>identifierType</i> zapíšeme OTHER, zde může být cokoliv, například ACCESSION NUMBER apod.	RA
<face>		Úroveň audioObject má nejméně jednu úroveň face. Face obsahuje jeden nebo více kanálů (streams)	M
	<i>ID</i>	Identifikátor této části dat, který se bude používat při odkazování z jiných částí záznamů (z region) Př. Side_1	M
	<i>direction</i>	Povolené hodnoty: FRONT BACK FORWARD REVERSE	M

		A_PASS B_PASS C_PASS D_PASS NONE	
	<i>audioObjectRef</i>	Odkaz na ID audioObject	M
	<i>label</i>	Název logického audio celku, např. název strany desky, pokud nějaký je (Rusalka, Půlnoční....) nebo Side_A. Tento název se pak může zobrazit uživateli v sw aplikaci.	M
<timeline>		Element obsahující subelementy popisující délku nahrávky a čas, kdy začíná 1-1	O
<startTime>		Začátek nahrávky, např. 0, udává se ve frames 1-1	O
	<i>editRate</i>	Vzorkovací frekvence popisovaného audio souboru, v herzích, např. "96000"	M
<duration>		délka nahrávky, udává se ve frames Př. 55642500 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného audio souboru, v herzích, např. "96000"	O

<region>		Další úroveň popisovaného audio objektu. Jednotlivé části nahrávky (face), úroveň face má nejméně jednu úroveň region. Úroveň region může odpovídat jednotlivým stopám. 1-n	M
	<i>ID</i>	ID této úrovně se odvozuje od ID vyšší úrovně face. K ID úrovně face se přidá 0001, 0002 apod., dle počtu regionů.	O
	<i>formatRef</i>	Obsahem atributu je odkaz na atribut ID elementu formatRegion	O
	<i>faceRef</i>	Obsahem atributu je odkaz na ID atribut vyšší úrovně face	R
	<i>label</i>	Název této úrovně audio objektu, např. Region_1 nebo název písně, proslovu apod.	R
<timeRange>		Element obsahující subelementy popisující délku popisovaného regionu a čas, kdy začíná 1-1	M
<startTime>		Začátek regionu, např. 0, udává se ve frames 1-1	M

	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<duration>		Délka regionu ve frames. 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<numChannels>		Počet kanálů (audio streams) regionu Zapísanou hodnotou je číslo, př. „1“ 1-1	M
<conditionNote>		Tento element může být použit k popisu stavu popisovaného objektu. 0-n	O
<note>		text poznámky 1-1	M
<noteTimeRange>		Element se použije k zápisu času (začátku a trvání) pro výše uvedenou poznámku ke stavu objektu 0-1	O
<startTime>		Začátek části nahrávky, pro který je platná poznámka, např. 0, udává se ve frames	M

		1-1	
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<duration>		Délka po kterou je poznámky platná ve frames. 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<creationDate>		Čas vytvoření poznámky, zápis dle normy ISO 8601 1-1	M
<securityNote>		Obsahem elementu je informace (volný text) o obsahu citlivých údajů (telefonních čísel, jmen...) 0-n	O
<stream>		Jednotlivé kanály, tj. u stereo nahrávky budou 2 elementy stream u mono nahrávky jen jeden element stream 1-n	M
	<i>ID</i>	Odvozuje se od ID vyšší úrovně (regionu), k id regionu se přidá	M

		písmeno, např. u008_20162011_021_0001a	
	<i>label</i>	Pojmenování kanálu. např. „1“ nebo „Stream1“	M
	<i>faceRegionRef</i>	Obsahem atributu je odkaz na ID regionu.	M
<channelAssignment>		Obsahem elementu je popis pozice kanálu. 1-1	M
	<i>channelNum</i>	číselné označení popisovaného kanálu, př. „1“	M
	<i>leftRightPosition</i>	Číselné označení umístění kanálu na ose levá-pravá. Hodnoty např. „0.0“ (střed) nebo „100.0“ (vpravo) nebo „-100.0“ (vlevo)	O
	<i>frontRearPosition</i>	Číselné označení umístění kanálu na ose...	O
<conditionNote>		Poznámka o stavu popisovaného kanálu (stream) 0-n	O
<note>		Text poznámky 1-1	M
<noteTimeRange>		Časový rozsah části kanálu, které se poznámka týká 0-1	O

<startTime>		Začátek části nahrávky, pro který je platná poznámka, např. 0, udává se ve frames 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<duration>		Délka po kterou je poznámky platná ve frames. 1-1	M
	<i>editRate</i>	Vzorkovací frekvence popisovaného regionu, v herzích, např. "96000"	M
<creationDate>		Čas vytvoření poznámky, zápis dle normy ISO 8601 1-1	M
<formatList>		Obsahem elementu jsou další charakteristiky fonoválečku	M
<formatRegion>			M
	<i>ID</i>	Identifikátor	M
	<i>label</i>	označení	M
	<i>ownerRef</i>		

<speed>		Obsahem elementu je popis rychlosti přehrávání a souvisejících informací.	O
<speedDesignated>		Rychlost přehrávání, zapíšeme číslo, např. „78“	M
	<i>unit</i>	Jednotka rychlosti přehrávání, např. „REVOLUTIONS_PER_MINUTE“	M
<varispeedAdjustment>		Element pro zápis úprav rychlosti při přehrávání/nahrávání.	O
	<i>unit</i>	Jednotka rychlosti přehrávání, např. „rpm“	M
<speedNote>		Volný text obsahující doplňující informace rychlosti nahrávky	O
<grooveOrientation>		Výběr z hodnot: LATERAL, VERTICAL, STEREO (Pro fonoválečky hodnota VERTICAL)	O
<grooveWidth>		Šířka drážky	O
<min>		Udává nejmenší zjištěnou vzdálenost mezi stěnami drážky	M
	<i>unit</i>	Jednotka rozměru, výběr z hodnot: MICROMETRES, MILLIMETRES, CENTIMETRES, MILS, INCHES, FEET, THREADS_PER_INCH	M
<max>		Udává největší zjištěnou vzdálenost mezi stěnami drážky	M
	<i>unit</i>	Jednotka rozměru, výběr z hodnot: MICROMETRES, MILLIMETRES, CENTIMETRES, MILS, INCHES, FEET, THREADS_PER_INCH	M

<grooveCreationMethod>		Výběr z hodnot: DIREC_CUT nebo PRESS_MOULDED	O
<soundField>		Obsahem elementu je popis zvukového pole, možné hodnoty jsou: „MONO“ nebo „STEREO“ nebo „SURROUND“. (Pro fonoválčky hodnota MONO) 1-1	O

7.5.5 Technická metadata MIX

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

Obsah pole "Popis":

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

Obsah pole "Použití pro"

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

Element	Popis	Povinnost	Použití pro
<BasicDigitalObjectInformation>			
<ObjectIdentifier>	údaje o identifikátoru obrazového dokumentu, který je formátem MIX popsán; 0-n	R	MC, PS

<objectIdentifierType>	např. jméno souboru, nebo jiný identifikátor; 0-1	M	MC, PS
<objectIdentifierValue>	hodnota identifikátoru, např. 20110306_001.jp2 nebo urn:nbn:123456; 0-1	M	MC, PS
<fileSize>	velikost souboru 0-1	R	MC, PS
<FormatDesignation>	údaje o formátu obrazového souboru 0-1	M	MC, PS
<formatName>	název formátu, např. lze využít MIME types (Image/jp2 apod.) 0-1	M	MC, PS
<formatVersion>	verze formátu, např. 1.0 0-1	M	MC, PS
<byteOrder>	endianita, možnosti jsou little endian, middle (mix) endian a big endian 0-1	M	MC, PS
<Compression>	údaje o kompresi obrazového souboru 0-n	M	MC, PS
<compressionScheme>	informace o kompresním schématu, vyjádřeno číslem (např. 34712 je komprese JPEG2000) nebo slovy (např. JP2 Lossless) 0-1	M	MC, PS
<BasicImageInformation>	základní technické údaje o obrazovém dokumentu 0-1	M	MC, PS
<BasicImageCharacteristics>	0-1	M	MC, PS
<imageWidth>	šířka obrazu v pixelech, např. 3987 0-1	M	MC, PS
<imageHeight>	výška obrazu v pixelech, např. 2345 0-1	M	MC, PS
<PhotometricInterpretation>	photometrická interpretace 0-1	M	MC, PS
<colorSpace>	barevný prostor, např. RGB 0-1	M	MC, PS
<ColorProfile>	údaje o barevném profilu 0-1 povinné pro dokumenty, kde je nutno uchovat přesnou reprezentaci barvy původního dokumentu a používá se ICC profil)	MA	MC, PS

<iccProfile>	ICC profil 0-1	M	MC, PS
<iccProfileName>	jméno profilu, např. sRGB, Adobe RGB aj. 0-1	M	MC, PS
<iccProfileVersion>	verze profilu, např. sRGB IEC61966-2.1 0-1	M	MC, PS
<iccProfileURL>	odkaz na profil, např. www.profilny.cz/sRGB_v4_ICC_pref.icc ; 0-1	R	MC, PS
<SpecialFormatCharacteristics>	speciální technické údaje o obrazovém dokumentu, povinné použití pro formát JPEG2000 0-1	MA	MC
<JPEG2000>	0-1	M	MC
<CodecCompliance>	údaje o kodeku 0-1	M	MC
<codec>	název kodeku, např. Kakadu, LuraWave aj. 0-1	M	MC
<codecVersion>	verze kodeku, např. 3.1 0-1	M	MC
<codestreamProfile>	popis codestream profilu JPEG2000, např. P0 a P1 (viz ISO/IEC 15444-4); 0-1	M	MC
<complianceClass>	specifikace největší výšky, šířky a počtu komponentů, které dekodér dokáže dekodovat, lze použít hodnoty C0, C1 a C2; 0-1	M	MC
<EncodingOptions>	obsahuje informace o kodování JPEG2000 0-1	M	MC
<Tiles>	popis pixelové velikosti dlaždic formátu JPEG2000 0-1	M	MC
<tileWidth>	šířka dlaždice, např. 128 0-1	M	MC
<tileHeight>	výška dlaždice, např. 128 0-1	M	MC
<qualityLayers>	číselná hodnota počtu vrstev, do kterých byl JPEG2000 rozdělen, např. 12 0-1	M	MC

<resolutionLevels>	popis počtu nižších rozlišení, které lze z obrazu získat, např. 6 0-1	M	MC
<ImageCaptureMetadata>	popis procesu skenování, je důležité vyplnit, protože tyto údaje nelze zjistit z finálního master/archivního souboru 0-1	M	PS
<SourceInformation>	informace o předloze 0-1	R	PS
<sourceType>	Book, Newspaper aj.; nutno používat kontrolovaný slovník 0-1	M	PS
<SourceID>	identifikátor předlohy 0-n	R	PS
<sourceIDType>	typ identifikátoru, např. čČNB, URN:NBN 0-1	M	PS
<sourceIDValue>	vlastní hodnota identifikátoru 0-1 povinné	M	PS
<GeneralCaptureInformation>	základní údaje o skenování 0-1	M	PS
<dateTimeCreated>	údaj o datu a čase skenování, např. 2009-01-03T08:25:28; zapsat v ISO 8601 na úroveň vteřin 0-1	M	PS
<imageProducer>	entita provádějící skenování, např. The National Library of the Czech Republic, osoba apod. 0-1	M	PS
<captureDevice>	typ skenovacího zařízení použít jedno z hodnot: <ul style="list-style-type: none"> - reflection print scanner (nejčrelection printtyp zaříp zat - transmission scanner - digital still camera - still from video 0-1	MA	PS
<ScannerCapture>	údaje o skeneru 0-1	M	PS
<scannerManufacturer>	výrobce skeneru, např. 4DigitalBooks, Treventus, Zeutschel 0-1	M	PS
<ScannerModel>	údaje o konkrétním typu skeneru 0-1	M	PS

<scannerModelName>	jméno modelové řady skeneru, např. DL 0-1	M	PS
<scannerModelNumber>	číslo/označení modelu, např. 3000 0-1	M	PS
<scannerModelSerialNo>	výrobní číslo skeneru, např. E4R0003649 0-1	M	PS
<MaximumOpticalResolution>	údaje o maximálním optickém rozlišení skeneru 0-1	M	PS
<xOpticalResolution>	optické rozlišení na ose x, pouze číselné vyjádření např. 300 0-1	M	PS
<yOpticalResolution>	optické rozlišení na ose y, pouze číselné vyjádření např. 300 0-1	M	PS
<opticalResolutionUnit>	jednotka optického rozlišení, jedna z hodnot: no absolute unit; in.; cm 0-1	M	PS
<scannerSensor>	popis typu snímacího senzoru skenovacího zařízení, jedna z hodnot: undefined; MonochromeLinear; ColorTriLinear; ColorSequentialLinear; MonochromeArea; OneChipColourArea; TwoChipColorArea; ThreeChipColorArea; ColorSequentialArea; 0-1	M	PS
<ScanningSystemSoftware>	údaje o softwaru skenovacího zařízení 0-1	M	PS
<scanningSoftwareName>	název softwaru, např. Copinet 0-1	M	PS
<scanningSoftwareVersionNo>	číslo verze softwaru, např. 3.7 0-1	M	PS
<DigitalCameraCapture>	údaje o snímacím zařízení (fotoaparát) 0-1 povinné, pokud je používán fotoaparát a není používán skener	MA	PS
<digitalCameraManufacturer>	výrobce fotoaparátu, např. Canon 0-1	M	PS
<DigitalCameraModel>	popis modelu fotoaparátu 0-1	M	PS
<digitalCameraModelName>	název modelové řady, např. EOS 0-1	M	PS

<digitalCameraModelNumber>	označení modelu fotoaparátu, např. 1000D 0-1	M	PS
<digitalCameraModelSerialNo>	výrobní číslo přístroje, např. E12345 0-1	M	PS
<camerarSensor>	typ senzoru fotoaparátu, např. matrix aj. 0-1	M	PS
<CameraCaptureSettings>	údaje o nastavení fotoaparátu použitého ke snímání předloh 0-1	M	PS
<ImageData>	v rámci tohoto kontejnerového elementu budou použity následující sub-elementy: fNumber exposureTime isoSpeedRatings shutterSpeedValue apertureValue brightnessValue exposureBiasValue maxApertureValue subjectDistance meteringMode lightSource flash focalLength backLight exposureIndex sensingMethod cfaPattern autoFocus PrintAspectRatio všechny hodnoty budou přebrány v případě použití fotoaparátu z údajů Exif	M	PS
<orientation>	popis orientace obrazu tak, jak je uložen vzhledem k jeho řádkům a sloupcům, např. normal*; normal, image flipper; normal, rotated 180°; unknown apod. 0-1	M	PS
<ImageAssessmentMetadata>	informace o digitálním obrazu pro jeho hodnocení a využití z hlediska dlouhodobé ochrany apod. 0-1	M	MC, PS
<SpatialMetrics>	rozměry obrázku, 2 rozměrná projekce objektů tak jak ji „vidí“ snímací zařízení 0-1	M	MC, PS

<samplingFrequencyPlane>	popis základní roviny, např. object plane (pro přímo ze předlohy digitalizované dokumenty), source object plane (pro digitalizaci mikrofilmů), camera/scanner focal plane (indikace sampl. frekvence fyzického senzoru); 0-1	R	MC, PS
<samplingFrequencyUnit>	jednotka měření sampl. frekvence hodnoty: "no absolute unit of measurement"; "in."; "cm"; 0-1	M	MC, PS
<xSamplingFrequency>	údaje o počtu pixelů na jednotku samplovací frekvence pro šířku obrázku 0-1 povinné, pokud hodnota samplingFrequencyUnit je "in." nebo "cm"	MA	MC, PS
<numerator>	čítatel, číselné vyjádření, např. 300 0-1	M	MC, PS
<denominator>	jmenovatel, číselné vyjádření např. 1 0-1	M	MC, PS
<ySamplingFrequency>	údaje o počtu pixelů na jednotku samplovací frekvence pro výšku obrázku 0-1 povinné, pokud hodnota samplingFrequencyUnit je "in." nebo "cm"	MA	MC, PS
<numerator>	čítatel, číselné vyjádření, např. 300 0-1	M	MC, PS
<denominator>	jmenovatel, číselné vyjádření např. 1 0-1	M	MC, PS
<ImageColorEncoding>	doplňující údaje o barvě obrazu 0-1	M	MC, PS
<BitsPerSample>	počet bitů na kanál 0-1	M	MC, PS
<bitsPerSampleValue>	hodnota počtu bitů, např. 8, 1, 4 nebo 8,8,8 apod. 0-n POZOR – pro každou hodnotu je nutno element opakovat, tj. např. 3x element <bitsPerSampleValue> s hodnotou 8 <mix:BitsPerSample> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> </mix:BitsPerSample>	M	MC, PS

<bitsPerSampleUnit>	specifikace jednotky, např. "integer" nebo "floating point" 0-1	R	MC, PS
<samplesPerPixel>	počet barevných komponentů na pixel, např. 1, 3, 4 0-1	M	MC, PS
<TargetData>	informace o kalibračních tabulkách 0-1 povinné pro obrazy, kde se dělá kontrola oproti kalibrační tabulce	MA	MC
<targetType>	typ kalibrační tabulky; 0= external (kalibrační tabulka se neobjeví na dig. obraze, je to oddělený dig. soubor); 1= internal (tabulka je naskenována spolu s přelohou a objeví se na dig. obraze); 0-n	M	MC
<targetID>	údaje o původu kalibrační tabulky 0-n	M	MC
<targetManufacturer>	výrobce/původce kalibrační tabulky, např. Eastman Kodak nebo NK ČR, oddělení kontroly kvality apod. 0-1	M	MC
<targetName>	název kalibrační tabulky, např. ColorChecker, MicrofilmScanTarget aj. 0-1	M	MC
<targetNo>	číslo nebo verze kalibrační tabulky 0-1	M	MC
<targetMedia>	údaj o tom, na jakém médiu je kalibrační tabulka, např. film, paper aj. 0-1	R	MC
<externalTarget>	údaje o externí kalibrační tabulce; např. link na 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
<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
<ImageProcessing>	údaje o zpracování obrazového souboru 0-n	M	MC

<dateTimeProcessed>	2009-01-04T15:12:06; zapsat v ISO 8601 na úroveň vteřin 0-1	M	MC
<sourceData>	odkaz na původní zdrojová data, ze kterých byl vytvořen finální obrazový soubor; může to být např. URL nebo cesta do složky s původním skenem včetně názvu souboru; 0-1	M	MC
<processingAgency>	The National Library of the Czech Republic 0-n	R	MC

7.6 METS část <fileSec>

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

7.6.1 <fileSec> hlavního záznamu METS

file group

- pro zvuk, obraz 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 původní zvukové soubory SA, archivní kopie zvukových souborů, uživatelské kopie zvukových souborů, archivní kopie obrazů, uživatelské kopie obrazů, další pro ALTO XML, OCR.TXT soubory a další pro METS záznamy s technickými metadaty (AMD_METS.xml)

<fileGrp> pro původní zvukové soubory SA - atributy ID="SA_AUDIOGRP" USE="source"

- o každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - kopíruje název souboru bez přípony
 - MIMETYPE - hodnota audio/wav
 - 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 link na zvukový soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku). Tj:
 - LOCTYPE - typ odkazu na soubor, typicky "URL"
 - xlink:href - odkaz na soubor

<fileGrp> pro zvukové soubory MCA - atributy ID="MC_AUDIOGRP" USE="master"

- každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID – kopíruje název souboru bez přípony
 - MIMETYPE – audio/wav
 - 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 zvukový soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<fileGrp> pro zvukové soubory UCA - atributy ID="UC_AUDIOGRP" USE="user"

- každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - kopíruje název souboru bez přípony
 - MIMETYPE - audio/mp3
 - 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 zvukový 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 bude mít následující atributy:

- i. ID="TECHMDGRP" USE="technical metadata"
- každý METS xml soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - identifikátor souboru amd_mets.xml jak je používán v METS záznamu; 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).

7.6.2 <fileSec> vedlejšího záznamu amd_mets.xml

- <fileSec> ve vedleším METS záznamu AMD_METS.xml bude obsahovat jeden element <fileGrp> s vnořenými elementy <file> pro každou reprezentaci zvukové stopy/stránky, tj. SA, MC pro zvuk, MC pro obraz, ALTO XML a OCR.TXT
- atributy jednotlivých <file> elementů odpovídají atributům pro jednotlivé typy dokumentů uvedených výše pro <fileSec> hlavního METS záznamu
- navíc pro SA a MC pro zvuk, 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

7.7 METS část <structMap> - Strukturální metadata

7.7.1 <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 dokument na úroveň skladby a obrazové přílohy
- vazby mezi konkrétním fonoválečkem a skladbami či konkrétní přílohou a obrázky jsou popsány v METS sekci <structLink>
- strukturální mapa logická i fyzická bude v hlavním záznamu main_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 zvukový dokument; v něm jsou dále vnořeny další elementy <div> druhé úrovně, které odkazují na všechny soubory patřící do zvukového dokumentu (zvukové stopy, obrázky, alto, ocr)
- <div> první úrovně obsahuje atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_0000)
 - LABEL - název zvukového dokumentu (nejvyšší úroveň)
 - TYPE - sound collection
 - DMDID - odkazuje na část bibliografických metadat náležejících ke zvukovému dokumentu
 - ADMID - identifikátor amdsec, ve kterém jsou obsažena (nepovinná) autorsko-právní metadata
- <div> druhé úrovně pro zvuk obsahuje tyto atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_AUDIO_0000)
 - TYPE - typ objektu, tj. sound
 - ORDER - reálné pořadí skladby na desce
 - ORDERLABEL - pořadové číslo skladby na fonoválečku, případně v jejich kolekci
 - ADMID - identifikátor amdsec, ve kterém jsou obsažena (nepovinná) autorsko-právní metadata
 - <div> druhé úrovně obsahuje vždy subelement <fptr>, v kterém je odkaz na soubory související se stopou (tj. SA, MCA, UCA, AMD_METS.xml)
 - o link na soubor má v elementu <fptr> atribut FILEID, který obsahuje ID souboru
- <div> druhé úrovně se zopakuje i pro obraz, bude mít atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_PAGE_0000)
 - TYPE - typ stránky (cover, imgdisc)
 - ORDER - pořadí stránky v příloze
 - ORDERLABEL - pořadové číslo stránky, jak je na ní vytištěno
 - ADMID - identifikátor amdsec, ve kterém jsou obsažena (nepovinná) autorsko-právní metadata
 - element <fptr> s linkem na konkrétní soubory (SA, MCA, UCA, MC, UC, OCR.TXT, ALTO XML, AMD_METS.xml)
 - o link na soubor má v elementu <fptr> atribut FILEID, který obsahuje ID souboru

- o 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"

7.7.1.2 Vyjádření logické strukturální mapy

7.7.1.2.1 Vyjádření logické struktury, s vazbou na ALTO bloky

- bude mít následující atributy: LABEL="logical_structure" TYPE="logical"
- <div> musí obsahovat identifikátor na příslušný blok metadat, př. odkaz na úroveň zvukového dokumentu DMDID="MODS_COLLECTION_0001"
- obrazové přílohy budou zanořeny přímo pod nejvyšší úroveň, tedy zvukový dokument
- element <div> první úrovně – zvukový dokument (soundcollection)
- element <div> druhé úrovně - obrazové informace (supplement, a rozlišit cover a imgdisc)
- element <div> druhé úrovně - skladba (soundrecording)
- výčet stop na jednom nosiči a výčet obrázků v příloze bude uveden v části <structLink>

<div> type	Atributy	Popis	Povinnost
SOUNDCOLLECTION	ID LABEL TYPE DMDID	<div> obsahuje údaje o nadřazeném zvukovém dokumentu ID - identifikátor <div>, hodnota např. SOUNDCOLLECTION_0001 LABEL - název celé kolekce TYPE - hodnota SOUNDCOLLECTION DMDID - obsahuje identifikátor popisné části MODS ke zvukovému dokumentu	M
SOUNDRECORDING	ID LABEL TYPE DMDID	<div> obsahuje údaje o skladbě ID - identifikátor <div>, hodnota např. SOUNDRECORDING_0001 LABEL - název skladby (číslo) TYPE - hodnota SOUNDRECORDING DMDID - obsahuje identifikátor popisné části MODS k části skladby	M

SUPPLEMENT	ID	<div> obsahuje údaje o obrazové příloze	M
	LABEL	ID - identifikátor <div>, hodnota např. SUPPL_0001	
	TYPE	LABEL - název přílohy - hodnoty COVER, IMGDISC	
	DMDID	TYPE - hodnota SUPPL	
		DMDID - obsahuje identifikátor popisné části MODS k příloze	

7.7.2 <structMap> vedlejšího záznamu METS (AMD_METS.xml)

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

METS část <structLink>

Bude ukazovat vztah mezi fyzickou a logickou strukturou, tedy vztah mezi jednotlivými zvukovými stopami na použitých úrovních a přílohou.

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

Př.:

```
<mets:structLink>
```

```
<mets:smLink xlink:from="SOUNDRECORDING_0001" xlink:to="DIV_AUDIO_0001"/>
```

```
<mets:smLink xlink:from="SOUNDRECORDING_0002" xlink:to="DIV_AUDIO_0002"/>
```

```
<mets:smLink xlink:from="SOUNDRECORDING_0003" xlink:to="DIV_AUDIO_0003"/>
```

```
<mets:smLink xlink:from="SOUNDRECORDING_0004" xlink:to="DIV_AUDIO_0004"/>
```

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

```
<mets:smLink xlink:from="SUPPL_0002" xlink:to="DIV_PAGE_0002"/>
```

```
<mets:smLink xlink:from="SUPPL_0003" xlink:to="DIV_PAGE_0003"/>
```

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

7.8 OCR (ALTO XML a TXT OCR)

- bude použita poslední verze formátu ALTO XML aktuální v době implementace, nebo verze předchozí (prosinec 2010 verze 2 – viz <http://www.loc.gov/standards/alto/>)
- níže uvedená specifikace neobsahuje všechny elementy a atributy formátu ALTO XML, obsahuje pouze ty, které jsou pro tuto konkrétní specifikaci relevantní – každý uvedený element má vyjádřenou míru relevance výrazy: povinné, doporučené a nepovinné
- elementy a atributy, které v této specifikaci nejsou uvedeny, nepovažujeme pro účely specifikace za důležité
- ALTO XML i OCR TXT vzniknou pro všechny obrazové soubory náležející k jedné intelektuální entitě (zvukovému dokumentu) 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, orientací a natočením obrazu, který byl použit pro vytvoření OCR
- OCR TXT bude vznikat z hotových ALTO XML během procesu digitalizace
- ALTO XML se bude vytvářet pouze pro dokumenty s určitou hranicí kvality OCR
- jméno OCR souboru musí odpovídat jménu obrazového souboru, ke kterému náleží; např. `pr_0007.jp2` a `al_0007.xml` nebo např. `123456_006_alto.xml` a `123456_006_archiv.jp2`
- kódování ALTO XML i TXT OCR musí být v UTF-8
- souřadnice pozic (HPOS, VPOS, WIDTH, HEIGHT) musí být vyjádřeny v pixelech
- v této specifikaci ALTO XML se počítá s OCR i pro text mimo tzv. textové "zrcadlo", tj. mimo hlavní text, jako jsou např. čísla stránek, běžící nadpisy a jiné části vyskytující se na okrajích stránky (top, left, top a bottom margin)
 - elementy `topMargin`, `leftMargin`, `rightMargin`, `bottomMargin` budou obsahovat elementy `<TextBlock>`, pro které platí stejná pravidla, jako pro element `<textBlock>` pro hlavní text stránky
 - pozor: údaje z OCR mimo hlavní text stránky by neměly být vyhledávatelné v aplikaci zpřístupnění, docházelo by ke zmatení uživatele a výsledků (např. při hledání titulu kapitoly by byly zobrazeny výsledky pro každou stránku, která obsahuje běžící nadpis apod.)
- pokud je na konci věty dělicí znaménko, ALTO XML i OCR TXT musí obsahovat oba fragmenty slova s dělícím 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

nejdou 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
 - 1-1 element je povinný, neopakovatelný
 - 0-n element je nepovinný, opakovatelný

Element	Atribut	Popis	Povinnost
<code><Description></code>			

<MeasurementUnit>		měřicí jednotka pro souřadnice v ALTO XML; možné hodnoty – dpi, pixel, inch1200 a mm10); inch1200 = 1/1200 inche; doporučené plnění je „mm10“ nebo „pixel“; 1-1	M
<sourceImageInformation>		informace o obrazovém souboru, ze kterého vzniklo ALTO XML; 1-1	MA
<fileName>		jméno obrazového souboru, ze kterého bylo ALTO XML vytvářeno; ideálně i s filesystem cestou jeho uložení; např. nlaImageSeq-33386-b.tif//produkce/OCR/digibok_XY/XY_011.tif 1-1	M
<fileIdentifier>		jedinečný identifikátor obrazového souboru; 0-n	R
<OCRProcessing>	ID	popis procesu vzniku OCR; 1-n ----- ID OCR procesu, např. <OCRProcessing ID="OCRPROCES_1">; povinné	M

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

<processingAgency>		jméno nebo kód instituce, např. NK CZ doporučujeme použít kontrolovaný slovník hodnot; 0-1	M
<processingSoftware>		popis SW, který dělal vlastní OCR; 1- 1	M
<softwareCreator>		výrobce softwaru - např. ABBYY, Russia; 1- 1	M
<softwareName>		jméno softwaru - např. FineReader; 1- 1	M
<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
<TextStyle>	ID FONTSTYLE FONTFAMIL Y 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

<ParagraphStyle>	ID ALIG N	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
<Page>	ID ACCURACY POSITION QUALITY PHYSICAL_IMG_ NR HEIGHT WIDTH PC	element popisující jednu stránku dokumentu; 1-n ----- ID – vygenerovaný identifikátor stránky, např. PAGE1, nebo P1 apod.; povinné ACCURACY – procentuální odhad přesnosti OCR (0-100); doporučené POSITION – pozice stránky; hodnoty k plnění: Left, Right, Foldout, Single, Cover; nepovinné QUALITY – krátký údaj o kvalitě předlohy stránky; hodnoty k plnění: OK, Missing, Missing in original, Damaged, Retained, Target, As in original; nepovinné PHYSICAL_IMG_NR - fyzické (pořadové) číslo stránky v dokumentu; vyjádřeno číslem, např. 1,2,3 apod.; povinné	M

		<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>	
<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
<TextBlock>		stejně plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
<LeftMargin>	ID HPOS VPOS WIDTH HEIGHT	<p>levý okraj – prostor mezi levým okrajem stránky a textem; 0-1</p> <p>-----</p> <p>ID: unikátní ID pro element LeftMargin, např. P1_LM0001 (page 1, leftMargin0001);</p>	M

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

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

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

<p><TextLine></p>	<p>ID STYLEREFS HPOS VPOS WIDTH HEIGHT</p>	<p>popis jedné řádky textu v rámci textového bloku; 1-n nutný alespoň jeden výskyt v rámci textového bloku</p> <p>-----</p> <p>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>	<p>M</p>
<p><String></p>	<p>ID CONTENT HEIGHT WIDTH HPOS VPOS CC WC</p> <p>V případě dělení slov také: SUBS_TYPE SUBS-CONTENT</p>	<p>řetězec znaků – vlastní obsah OCR; znaky tvoří jednotlivá slova a více tagů <String> větu <TextLine>; 1-n v rámci <TextLine></p> <p>-----</p> <p>ID obsahuje unikátní sekvenční číslo řetězce na stránce, např. "P3_ST0001" (strana 3, řetězec 1); povinné</p> <p>CONTENT – ukládá vlastní řetězec znaků (slovo); povinné</p> <p>HPOS: horizontální pozice řetězce;</p>	<p>M</p>

		<p>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;</p>	
--	--	--	--

		<p>HypPart2 se vyskytuje u následujícího tagu <CONTENT> v následující větě (stringu), který obsahuje druhou část rozděleného slova/řetězce;</p> <p>Abbreviation – typ substituce používaný při rozepisování zkratk v textu na jejich plný text; při dělení slov v textu HypPart1 a HypPart2 povinné, abbreviation nepovinné</p>	
<ALTERNATIVE>		<p>alternativní hodnota OCR řetězce pro jednotlivá slova;</p> <p>0-n</p> <p>lze použít v případě nejistoty rozpoznání řetězce;</p>	O
<HYP>	<p>CONTENT</p> <p>WIDTH</p> <p>HPOS</p> <p>VPOS</p>	<p>zápis znaku rozdělovníku slov</p> <p>0-1 pro jeden výskyt <TextLine>; vždy pro poslední <String>;</p> <p>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 „-“;</p> <p>povinné</p> <p>WIDTH – šířka dělicího znaku;</p> <p>doporučené</p> <p>HPOS: horizontální pozice dělicího znaku;</p> <p>doporučené</p> <p>VPOS: vertikální pozice dělicího znaku;</p> <p>doporučené</p>	MA

<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
<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 grafík</p> <p>-----</p> <p>ID: unikátní ID komponovaný blok, např. P6_CB0001 (stránka 6, komponovaný blok 0001); povinné</p> <p>TYPE – označení typu komponovaného bloku; nutné používat kontrolovaný slovník (illustration, Advertisement, apod.); povinné</p>	MA

		<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>	
<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
<Polygon>	POINTS	<p>popis tvaru víceúhelníku; 1-1 ----- POINTS – vyjádření jednotlivých bodů víceúhelníku povinné</p>	M

<TextBlock>	ID STYLEREFS HPOS VPOS WIDTH HEIGHT	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 STYLEREFS: reference na ID definice formátování textových bloků /alto/Styles/ParagraphStyle; povinné HPOS: horizontální pozice bloku; povinné VPOS: vertikální pozice bloku; povinné WIDTH – šířka textového bloku; povinné HEIGHT – výška textového bloku; povinné	MA
<TextLine>		/alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock/TextLine a ostatní elementy v rámci /alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock mají stejná pravidla a výskyty jako jako ve vrchním elementu /alto/Layout/Page/PrintSpace/TextBlock	

<GraphicalElement>	ID HPOS VPOS WIDTH HEIGHT	popis grafického tvaru; v případě využití v rámci /alto/Layout/Page/PrintSpace/ComposedBlock označuje rozměry tvaru v rámci něhož je tabulka, ilustrace, reklama apod.; 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.; ----- ID – identifikátor grafického tvaru; povinné HEIGHT – výška grafického tvaru; povinné WIDTH – šířka grafického tvaru; povinné HPOS – horizontální pozice grafického tvaru; povinné VPOS – vertikální pozice grafického tvaru; povinné	MA
--------------------	---------------------------------------	--	----

8 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). 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

Část METS pro autorsko-právní metadata

element	atributy	popis	Povinnost
<amdSec>		element obsahující autorsko-právní metadata ve formátu ContainerMD	M
	ID	Použití jak u hlavního METS, tak vedlejšího záznamu amd_mets ID – identifikátor konkrétní části <amdSec>, např. pro celý zvukový dokument hodnota "AMD_SOUNDCOLLECTION_0001", pro 1 skladbu hodnota "AMD_SOUNDRECORDING_0001"	M

<rightsMD>		element pro typ autorsko-právních metadat	M
	ID	ID pro část <rightsMD>: např. "RIGHTS_VOLUME", RIGHTS_0001; ...	M
<mdWrap>	MDTYPE	MDTYPE – hodnota "OTHER"	M
	OTHERMDTYPE	OTHERMDTYPE – hodnota "CopyrightMD"	M
	MIMETYPE	MIMETYPE – hodnota "text/xml"	M
<xmlData>			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 list^[3]) (viz příklad)
2. *year.type*: elementy: <year.copyright> <year.renewal> <year.publication> <year.creation> <year.birth> <year.death> - do atributu se zapisuje přesnost uvedeného roku. Možné hodnoty *exact* (rok je známý), *approximate* (odhadovaný rok na základě jiných zdrojů), *unknown* (rok není znám a není možné jej odhadnout/dohledat)

Element	Popis	Povinnost
---------	-------	-----------

<copyright>	<p>Povinný kořenový element. Označuje status dokumentu. 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) • 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
<creation>	<p>Informace o vytvoření dokumentu.</p> <p>0-1</p>	O
<year.creation>	<p>Rok vytvoření dokumentu ve formátu YYYY.</p> <p>0-1</p>	RA
<country.creation>	<p>Země, ve které byl dokument vytvořen: vyjádřeno iso kódem (viz Nepovinné atributy) př.: <country.creation iso.code="cz">Česká republika</country.creation></p> <p>0-1</p>	RA
<creator>	<p>Informace o tvůrci dokumentu.</p> <p>0-1</p>	MA
<creator.corporate>	<p>Název instituce, pokud je autorem dokumentu.</p> <p>0-n</p>	MA

<creator.person >	Informace o autorovi dokumentu. 0-n	MA
<name>	Jméno a příjmení autora. 0-1	MA
<year.birth>	Datum narození autora. Formát YYYY. 0-1	MA
<year.death>	Datum úmrtí autora. Formát YYYY. 0-1	MA
<note>	Doplňující informace k tvůrci dokumentu. 0-n	O
<publication>	Informace o vydání dokumentu 0-1	MA
<country.public ation>	Země vydání – vyjádřeno iso kódem (viz Nepovinné atributy), př: <country.publication iso.code="cz">Česká republika</copyrightMD.country.publication> 0-1	R
<publisher>	Nakladatel. 0-1	R
<year.publicatio n>	Rok vydání. Formát YYYY. 0-1	MA
<year.copyright>	Rok copyrightu. Formát YYYY. 0-1	RA
<year.renewal>	Rok případného obnovení copyrightu. Formát YYYY. 0-1	RA

<note>	Doplňující informace k vydání dokumentu. 0-n	O
<rights.holder>	Informace o držiteli práv. Používá se v případě, že práva drží někdo jiný než autor nebo nakladatel. Např. občanské sdružení autorů apod. 0-1	O
<contact>	Kontakt na držitele práv. 0-n	O
<name>	Jméno nebo název držitele práv. 0-1	O
<note>	Doplňující informace o držiteli práv. 0-n	O
<notice>	Do pole se vkládá oznámení o copyrightu, tak jak je uvedeno v dokumentu. 0-1	O
<general.note>	Pole slouží k vložení jakékoliv další informace, která je relevantní vzhledem k autorským právům dokumentu a nemohla být zařazena do žádného jiného pole. 0-n	O

[1] <http://tools.ietf.org/html/rfc4234>

[2] toto pořadí nevyjadřuje logickou strukturu dokumentu, ta je popsána dále v DMF.

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