

# Zvukové dokumenty ve fondech paměťových institucí v kontextu dlouhodobého uchovávání v ČR. Přípravná studie NK ČR k možnosti dlouhodobého uložení digitalizovaných dat<sup>1</sup>

Audio documents in the collections of memory institutions in context of long term preservation in the Czech Republic. Feasibility study by National Library of the Czech Republic about the possibility of long term preservation of digitized items

*Mgr. Natalie Ostráková, Filip Šír, Dis. / Národní knihovna České republiky (National Library of the Czech Republic), Klementinum 190, 110 00 Praha 1, Česká republika, Národní muzeum (National Museum), Václavské nám. 68, 115 79 Praha 1, Česká republika*

## Resumé:

V uplynulém období byla v České republice věnována velká pozornost digitalizaci tištěných materiálů a dlouhodobému uchovávání dokumentů vzniklých digitalizací. Nyní přichází na řadu další, do té doby téměř opomíjená, cenná část našeho kulturního dědictví – zvukové dokumenty. Ve spolupráci s Moravskou zemskou knihovnou se Národní knihovna ČR v minulém roce začala věnovat dlouhodobému uchovávání zvukových dokumentů. Byla provedena analýza situace ve světě, prozkoumány standardy a doporučení mezinárodních organizací (ARSC<sup>2,3</sup>, IASA<sup>4,5</sup>, LOC<sup>6</sup> apod.) i praxe vybraných zahraničních institucí – knihoven, archivů, institucionálních repozitářů. Byly též testovány softwarové nástroje pro práci se souborovými formáty zvukových dokumentů. Na základě zjištěných údajů pak byly zvoleny souborové formáty pro archivaci a zpřístupnění zvukových dokumentů, metadatová schémata pro jejich popis a nástroje pro identifikaci, validaci a charakterizaci.

**Klíčová slova:** zvukové dokumenty, dlouhodobá ochrana digitálních dokumentů, dlouhodobá archivace, metadata, souborové formáty

## Summary:

A great attention has been paid in the Czech Republic in the recent past to the digitization of printed materials and a long term preservation of digitized documents. Now is the turn of another, until now almost neglected, valuable part of our cultural heritage – audio documents. In cooperation with the Moravian Library in Brno, the National Library of the Czech Republic started last year to look at long term preservation of audio documents. Analysis was made of the situation in the rest of the world and standards and recommendations of international organisations (ARSC<sup>2,3</sup>, IASA<sup>4,5</sup>, LOC<sup>6</sup> etc.) as well as the practice of selected foreign institutions – libraries, archives, institutional repositories

<sup>1</sup> Realizováno díky institucionálnímu výzkumu Národní knihovny České republiky financovanému Ministerstvem kultury ČR v rámci Dlouhodobého koncepčního rozvoje výzkumné organizace.

<sup>2</sup> ARSC Guide to Audio Preservation, <https://www.clir.org/pubs/reports/pub164>

<sup>3</sup> <http://www.arsc-audio.org/index.php>

<sup>4</sup> <http://www.iasa-web.org/>

<sup>5</sup> Guidelines on the Production and Preservation of digital Audio Objects, <http://www.iasa-web.org/tc04/audio-preservation>

<sup>6</sup> Library of Congress Recommended Formats Statement, <https://www.loc.gov/preservation/resources/rfs/>

were examined. Software tools for work with file formats of audio documents were also tested. On the basis of the data found, file formats were selected for archiving and accessibility to audio documents together with metadata schemes for their description and tools for identification, validation and characterization.

**Keywords:** audio documents, digital preservation, long term preservation, digital formats, metadata

Národní knihovna ČR (NK ČR) již několik let provozuje dlouhodobé digitální úložiště, které je však zatím určeno jen pro digitalizovaná periodika a monografické dokumenty. V rámci vývoje svého řešení pro dlouhodobé uchovávání se NK ČR nyní začala, ve spolupráci s Moravskou zemskou knihovnou, věnovat i otázce dlouhodobého uchovávání zvukových dokumentů. Tento článek představuje vstupní analýzu, která byla provedena jako podklad pro vypracování standardu pro dlouhodobé uložení a zpřístupnění zvukových dokumentů. Analýza se věnuje třem oblastem – souborovým formátům, metadatům a nástrojům pro zvukové dokumenty. Předmětem analýzy jsou současné postupy vybraných paměťových institucí při dlouhodobém uchovávání a zpřístupňování zvukových dokumentů (doporučované formáty, metadatová schémata a nástroje). Výstupem analýzy budou zásady, jak definovat národní standard pro uchovávání a zpřístupňování zvukových dokumentů pro české paměťové instituce. Základem standardu bude specifikace archivačního balíčku (AIP, tj. Archival Information Package) pro zvukové dokumenty, která bude v souladu s osvědčenými mezinárodními postupy. Z balíčku AIP bude následně vytvořena specifikace pro balíčky SIP (Submission Information Package) pro dodavatele dat. Specifikace balíčku SIP pro dodavatele umožní standardizaci digitalizace zvukových dokumentů s ohledem na jejich dlouhodobé uchovávání, čímž se usnadní jejich převod do balíčku AIP v dlouhodobém úložišti NK ČR.

## Výchozí situace v oblasti uchovávání zvukových dokumentů v České republice

Před uvedením samotné studie je potřeba popsat výchozí situaci, v níž se paměťové instituce, resp. knihovny v České republice, nyní nacházejí.

Národní zvukový archiv, instituce spravující zvukové dokumenty a koncepčně řešící výše zmíněnou problematiku, se v ČR bohužel nenachází. S tímto hlavním bodem výchozí situace je třeba počítat. Zatímco ve světě je typ národní fonotéky, archivu apod. obvyklý – např. British Library<sup>7</sup>, Library of Congress<sup>8</sup>, Bibliothèque nationale de France<sup>9</sup> – u nás se tuto ideu nepodařilo prosadit ani našim předchůdcům, ani současným generacím. Fakt neexistence národního zvukového archivu se nepříznivě promítá do celé řady aspektů. Vede např. k tomu, že se o tyto dokumenty instituce nestarají tak, aby byly schopny garantovat jejich evidenci, ochranu, digitalizaci a zejména zpřístupnění, resp. uchování pro budoucí generace. Dalším aspektem je to, že zvukové dokumenty, jakožto obvykle menší součást heterogenních fondů knihoven, archivů či muzeí, dosud nebyly a nejsou prioritou: stojí až na pomyslné druhé koleji za tištěnými dokumenty. Je třeba také konstatovat, že zvukové fondy dosud nebyly koncepčně a spolehlivě zmapovány, resp. zpracovány, a tedy je nelze ani zpřístupnit. To se negativně poděpisuje na dalších

<sup>7</sup> <http://sounds.bl.uk/>

<sup>8</sup> <http://www.loc.gov/jukebox/>

<sup>9</sup> <http://gallica.bnf.fr/>

procesech, které by měly přispět ke koncepčnímu řešení vedoucímu až k samotné digitalizaci, tedy k dlouhodobému uchování zvukových dokumentů. Celou tuto situaci ovlivňuje i fakt, že výše zmíněné procesy – katalogizace, digitalizace, zpřístupnění – jsou v případě zvukových dokumentů poněkud náročnější než u dokumentů textových či obrazových (Horová, Novotná, Šír 2016).

Řešením, které se za dané situace nabízí, je pověření některé instituce zřizované Ministerstvem kultury ČR (dále jen MK ČR), která by měla v určitou chvíli začít koncepčně koordinovat práci spjatou s danou problematikou a metodicky vést skupinu jedinců a institucí postupnými kroky k naplnění tří základních postupů – evidence, digitalizace, zpřístupnění.

Rozhodnutím MK ČR byla v roce 2015 pověřena vykonáváním koordinační a metodické role v oblasti zvukových dokumentů a rozvojem Virtuální národní fonotéky Moravská zemská knihovna v Brně (dále MZK) (Rozhodnutí ministra kultury 2015). V Konceptci MZK na léta 2015–2020 je již zohledněna nová situace: „*Moravská zemská knihovna bude organizovat a koordinovat spolupráci předních institucí a organizací, které uchovávají materiály národního zvukového dědictví s cílem metodicky a organizačně podpořit katalogizaci a digitalizaci zvukových sbírek u nás a budovat portál Virtuální národní fonotéka jako bránu zpřístupňující na jednom místě informace o zvukových dokumentech a tvořící základní nástroj pro jejich následnou digitalizaci.*“ (Konceptce Moravské zemské knihovny 2015). V ČR tedy existuje instituce, která má tuto problematiku koordinovat a metodicky napomáhat ke konkrétním krokům v celém procesu ochrany a zpřístupnění zvukových dokumentů.

Pro „uchopení“ celého komplexu těchto procesů (ochrany a evidence, následně digitalizace a dlouhodobého uchování) je nutné mít na vědomí základní principy práce se zvukovými či audiovizuálními dokumenty. Je zde potřeba připomenout nejen technické standardy digitalizace mezinárodních asociací ARSC (Association for Recorded Sound Collections)<sup>10</sup> či IASA (International Association of Sound and Audiovisual Archives)<sup>11</sup>, které budou podrobněji zkoumány později, ale především dokument, který by měl v našem prostředí zahájit realizaci projektů, změnit názory vlastníků sbírek či podpořit snahu nejen Národní knihovny ČR, ale i ostatních paměťových institucí. Vancouverští deklarace (Vancouver Declaration) – mezinárodní dokument UNESCO/UBC vzniknuvší na základě jednání konference Paměť světa v digitálním věku: Digitalizace a uchování (The Memory of the World in the Digital Age: Digitization and Preservation, UNESCO/UBC 2012) říká:

*V současné době se digitální informace vytrácí, protože její hodnota je podceňována, a to z důvodu absence právních a institucionálních rámců nebo proto, že kurátoři nemají znalosti, dovednosti a finanční prostředky.*

Pokud se pozorně podíváme na tuto větu, je zřejmé, že je třeba začít koncepčně pracovat a začít vzdělávat nejen kurátory, knihovníky, muzejníky, ale i vedení institucí spravujících tyto sbírky. V případě zvukových dokumentů je třeba mít povědomí o všech aspektech nastavení v rámci digitalizace, ale také vypracovat návrhy řešení pro adekvátní popis všech typů metadat.

Výsledky této mezinárodní konference z roku 2012 jsou veřejně známé, ale i tak je třeba připomenout základní pilíře, které je bezpodmínečně nutné brát v úvahu a apliko-

<sup>10</sup> <http://www.arsc-audio.org/>

<sup>11</sup> <http://www.iasa-web.org/>

vat na projekty všeho druhu i v oblasti paměťových institucí ČR. Více než 500 účastníků se dohodlo, že:

2. *v případě analogových dokumentů může digitalizace ochránit cenné dokumenty před znehodnocením minimalizací manipulace s nimi. V případě audiovizuálních dokumentů je digitalizace jediným prostředkem k zajištění jejich přežití;*

To ale znamená, že je třeba mít takové postupy, které by byly aplikovatelné v různých typech institucí či by bylo možné je využít při vzdělávání informačních pracovníků a odborníků zabývajících se těmito dokumenty.

6. *musí být vytvořeny vzdělávací a školicí programy pro informační profesionály, které musí být dostupné pro přípravu nebo přeškolení za účelem zavádění postupů digitalizace a uchovávání relevantních podkladů pro potřeby vlády a jejich občanů.*

## Zvukové dokumenty ve strategických a koncepčních materiálech MK ČR

Problematika zvukových dokumentů představovala do roku 2015 opomíjenou část kulturního dědictví: nefigurovala v žádném oficiálním strategickém dokumentu na národní úrovni, ani nebyla koncepčně řešena v rámci národních institucí. Teprve až díky aktivní podpoře Moravské zemské knihovny v Brně, zapojením se do Mezinárodní asociace IASA a otevřením tohoto tématu v Ústřední knihovnické radě se do Plánu implementace Státní kulturní politiky na léta 2015–2020 podařilo vložit samostatný oddíl věnující se dané problematice – „3.3.2 Uchování národního zvukového dědictví“ (Plán implementace Státní kulturní politiky, 2015, s. 47–48). Samotný oddíl obsahující níže uvedené body je v rámci této studie a následné práce na dlouhodobém uchovávání stěžejní:

„3.3.2 Uchování národního zvukového dědictví

3.3.2.A: Koncepce ochrany a zpřístupnění zvukových dokumentů jako významné součásti kulturního dědictví

3.3.2.B: Vybudování metodického a digitalizačního centra pro záchranu zvukových dokumentů v Moravské zemské knihovně v Brně ve spolupráci s Národním muzeem (České muzeum hudby)

3.3.2.C: Zajistit dlouhodobé úložiště pro uchování a zpřístupnění zvukových dat v rámci resortního LTP systému Ministerstva kultury

3.3.2.D: Provozovat a rozvíjet portál Virtuální národní fonotéka (role sektorového agregátora)<sup>12</sup>

Jak je vidět v bodech 3.3.2.B a 3.3.2.C, pro uchování zvukových dokumentů je zde zakotvena jak fáze metodiky a digitalizace, tak neméně důležitá součást – dlouhodobé úložiště. Ve zmíněném dokumentu je zakotvena podpora Národní knihovny ČR při řešení problému dlouhodobého uložení zvukových dokumentů, resp. zvukových formátů.

## Formáty pro dlouhodobé uchovávání zvukových dokumentů

Data jsou v počítačích ukládána v souborovém formátu, který určuje vnitřní strukturu a pravidla a postupy práce s nimi. Jednotlivé formáty se liší možnostmi komprese,

<sup>12</sup> <https://www.narodnifonoteka.cz/>

odolností vůči chybám, existencí, dostupností a otevřeností specifikace (např. standardu), rozšířením mezi uživateli apod. Existuje celá řada formátů pro uložení zvukových dokumentů, pro potřeby paměťových institucí jsou ale vhodné jen některé z nich. Zvláštní pozornost je třeba věnovat volbě archivačního formátu (formátu pro dlouhodobé uchovávání). Ten by měl ideálně splňovat následující podmínky (Library of Congress 2013; Rog, Wijk 2008):

1. být otevřeným formátem s veřejně dostupnou a podrobnou specifikací,
2. být dostatečně rozšířený, aby existoval zájem o jeho podporu i v budoucnosti a aby existovaly nástroje pro jeho migraci a validaci,
3. být transparentní – mít jednoduchou strukturu a být analyzovatelný existujícími nástroji (tj. nezašifrovaný a ideálně nekomprimovaný),
4. být sebedopisný – tj. s možností vložit metadata přímo do digitálního objektu (popisná, technická...), jež usnadní správu objektu,
5. být robustní, tj. odolný vůči chybám,
6. být nezávislý na konkrétním hardwaru, operačním systému, softwaru.

Výše uvedené vlastnosti archivačního formátu pro zvukové dokumenty splňuje nejlépe zvukový signál kódovaný metodou pulzně kódové modulace (PCM) a kontejnerové formáty WAVE<sup>13</sup>, BWF<sup>14</sup>, AIFF<sup>15</sup>. PCM je metoda převodu analogového zvukového signálu na signál digitální. Klíčovými parametry kódování zvuku jsou vzorkovací frekvence a bitová hloubka. Tyto parametry ovlivňují věrnost zachyceného zvuku (tj. výslednou kvalitu zvuku). Pro dlouhodobé uchovávání jsou aktuálně doporučovány hodnoty: 24bitová (a vyšší) hloubka a tzv. vzorkovací frekvence nejméně 48 kHz, ideálně 96 kHz a vyšší (International Association of Sound and Audiovisual Archives 2009; Brylawski 2015). Tyto hodnoty, jež jsou poměrně vysoko nad hranici lidského sluchu, umožňují zachytit informace, které sice člověk slyšet nemůže, ale zvukový dokument je přesto obsahuje<sup>16</sup> (CDP Digital audio working group 2006). Tyto vyšší hodnoty též mají umožnit posluchačům lépe vnímat prostor, hloubku a další vlastnosti zvuku (Consortium of Academic and Research Libraries in Illinois 2014). V některých případech jsou ale přípustné nižší hodnoty, například:

1. Nosiče CD, DAT<sup>17</sup> apod., které obsahují zvuková data o hodnotách 44,1 kHz a 16 bitů. Digitalizace při hodnotách 96/24 by neměla vliv na kvalitu výsledného zvuku a soubor by se jen zvětšil.
2. Některé zdroje uvádějí, že pro mluvené slovo je dostačující nižší vzorkovací frekvence<sup>18</sup>. Tak Konsorcium akademických a výzkumných knihoven v americkém státě Illinois pro digitalizaci mluveného slova bez hudebního doprovodu či dalších, např. zvířecích zvuků v pozadí, doporučuje vzorkovací frekvenci 48 kHz (Consortium of Academic and Research Libraries in Illinois 2014).

<sup>13</sup> Waveform Audio File Format, též WAVE nebo WAV, <http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/Docs/riffmci.pdf>, s. 56–65.

<sup>14</sup> Broadcast Wave Format, též BWF, <https://tech.ebu.ch/docs/tech/tech3285.pdf>

<sup>15</sup> Audio Interchange File Format, též AIFF, <http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/AIFF/Docs/AIFF-1.3.pdf>

<sup>16</sup> Například některé zvuky, které vydávají zvířata a elektronické instrumenty, vykazují frekvenci vyšší než 22,05 kHz, tj. nad hranici lidského sluchu.

<sup>17</sup> DAT = Digital audio tape, zvuk v digitální podobě uložený na páskovém nosiči.

<sup>18</sup> HAMILTON COLLEGE LIBRARY. Creating/Submitting Digital Audio. *Digital Collections-Hamilton College Library* [online]. Clinton, NY, 2013 [cit. 2016-12-02]. Dostupné z: <http://elib.hamilton.edu/digital-audio>.

Praxe jednotlivých institucí se může lišit. Například audiokazety a magnetofonové pásky digitalizují v Knihovně J. Willarda Marriotta Univerzity v Utahu za použití vzorkovací frekvence 44,1 kHz a bitové hloubky 24 (J. Willard Marriot Library 2015). V Knihovně Univerzity v Yorku pro tyto nosiče preferují hodnoty vyšší – ideálně bitovou hloubku 32 (minimálně 24) a vzorkovací frekvenci ideálně 96 kHz (minimálně 48 kHz) (York University Libraries 2017).

Formáty WAVE, BWF a AIFF jsou kontejnerovými formáty, což znamená, že kromě zvukových dat mohou nést i další objekty, například metadata. Nejčastěji nesou tyto kontejnery zvuková data kódovaná metodou PCM<sup>19</sup> (Florida center for library automation, c2008).

Formát WAVE (Waveform Audio File Format) vytvořily v roce 1991 společnosti Microsoft a IBM. Má jednoduchou strukturu a je jedním z nejjednodušších formátů pro uložení audia (Remo Software 2015), což jej činí ideálním pro dlouhodobé uchování. Je kompatibilní s operačními systémy Windows, OS X i Linux. V hudebním a vysílacím průmyslu je hojně rozšířený, což má za následek vysokou podporu v aplikacích a hardwarových zařízeních. Je tedy velmi pravděpodobná i určitá podpora formátu v budoucnosti.

Modifikací formátu WAVE (přidáním zvláštního prostoru určeného výhradně pro zápis metadat) vytvořila Evropská vysílací unie v roce 1997 formát BWF (Broadcast Wave Format). Tento formát se stal de facto standardem pro vysílací instituce a jeho užití se považuje za osvědčený postup pro ukládání digitálního audia (Whibley 2016). Je otevřený, široce rozšířený, transparentní, sebedokumentující, nezávislý na hardwaru a softwaru a není omezen patenty.

V roce 2008 iniciativa Florida Voices spadající pod Floridskou elektronickou knihovnu uvedla, že WAVE a BWF jsou nejužívanějšími archivačními formáty ve Spojených státech amerických. (Florida center for library automation 2008). A zdá se, že situace je stejná i v roce 2016 – formáty WAVE a BWF jsou podle Týmu pro digitální archivaci britské knihovny (Digital Preservation Team) stále nejrozšířenějšími formáty pro dlouhodobé uchování zvukových dokumentů (Whibley 2016). Formát BWF jako archivační formát doporučují také tři významné organizace – Mezinárodní asociace zvukových a audiovizuálních archivů (IASA) (International Association of Sound and Audiovisual Archives 2009), Kongresová knihovna (Library of Congress 2016) a asociace ARSC (Association for Recorded Sound Collections) (Brylawski 2015). Doporučení prezentovaná organizací IASA v publikaci „*TC-04 Guidelines on the Production and Preservation of Digital Audio Objects*“ (International Association of Sound and Audiovisual Archives 2009) se zdají být v současnosti nejužívanějšími pro dlouhodobé uchování zvukových dokumentů. Formáty WAVE a BWF jako archivační formáty aktuálně přijímají a doporučují například tyto instituce: Britská knihovna, *British Broadcasting Corporation* (BBC) (Whibley 2016), Moravská zemská knihovna (Šír, Žabička 2013), Knihovny a archivy Kanady (Library and Archives Canada 2015)<sup>20</sup>, digitální repozitář DataShare univerzity v Edinburghu (University of Edinburgh 2015)<sup>21</sup>, digitální repozitář Coloradské státní univerzity (Colorado State University 2015) (preferují ale WAVE)<sup>22</sup>, Knihovna J. Willarda

<sup>19</sup> Mohou ale nést i data s kompresí MP3 a mít pak koncovku souboru .wav a nikoliv .mp3.

<sup>20</sup> <http://www.bac-lac.gc.ca/eng/services/government-information-resources/guidelines/Pages/guidelines-file-formats-transferring-information-resources-enduring-value.aspx#p>

<sup>21</sup> [http://www.ed.ac.uk/files/atoms/files/recommended\\_file\\_formats-apr2015.pdf](http://www.ed.ac.uk/files/atoms/files/recommended_file_formats-apr2015.pdf)

<sup>22</sup> <http://lib.colostate.edu/repository/csu-digital-repository-preservation-format-support-policy>

Marriotta na univerzitě v Utahu<sup>23</sup>, institucionální repozitář univerzity v Minnesotě<sup>24</sup> (University of Minnesota c2014) apod.

Formát AIFF, jenž v roce 1988 vytvořila společnost Apple Computers, je sice hojně rozšířen mezi uživateli počítačů Macintosh, není ale tak široce podporován jako formát WAVE. Většina uživatelů OS Windows používá formát WAVE (Library of Congress 2014). Ve formátových politikách paměťových institucí se formát AIFF objevuje méně často než formát WAVE<sup>25</sup> (Rimkus et al. 2014). Jako preferovaný formát jej spolu s formátem WAVE uvádí například Státní knihovna Nový Jižní Wales (State Library New South Wales 2015).

V případě, že je organizace limitována úložnou kapacitou, lze zvuková data bezeztrátově komprimovat. Za tímto účelem lze využít bezeztrátový formát/kodek FLAC. Kodek FLAC komprimuje zvuková data bez jakékoli ztráty a výsledný zvuk je při poslechu stejně kvalitní jako zvuk ve formátu WAVE. Konverze souborů WAVE a BWF do formátu FLAC je teoreticky kdykoliv bezeztrátově vratná (Xiph.org Foundation 2014). Formát FLAC vyniká nad formáty BWF a WAVE v několika ohledech. Dokáže snížit velikost souboru až o dvě třetiny původního souboru. Navíc každý soubor FLAC obsahuje v hlavě souboru kontrolní součet (algoritmus MD5) původního souboru BWF/WAVE, čímž se usnadní kontrola integrity dat, pro kterou není nutné získávat informace z externího souboru s kontrolními součty (Rice 2013).

Formát FLAC je spolu s BWF preferovaným archivačním formátem Národního archivu Spojených států amerických, jenž dlouhodobě uchovává vládní a historické záznamy Spojených států amerických (The U.S. National Archives and Records Administration 2014). Další organizace, např. Bentleyova historická knihovna (Bentley historical Library 2016), digitální repozitář DataShare univerzity v Edinburghu (University of Edinburgh 2015), knihovny Univerzity York (York University Libraries 2015), britský datový archiv (UK Data Archive 2016), Tasmánský archiv (Tasmanian Archive and Heritage Office 2015), knihovny univerzity Binghamton (Binghamton University Libraries c2016) formát FLAC doporučují a přijímají spolu s formáty WAVE, BWF a/nebo AIFF. BBC používá formát WAVE, pro hodně velké soubory ale nevyklučuje použití formátu FLAC (Whibley 2016). Národní archiv Austrálie (National Archives of Australia) formát FLAC nyní zvažují jako svůj druhý archivační formát<sup>26</sup>, do nějž budou normalizovány dodané formáty (Crowe 2016). V nizozemské Národní knihovně se nyní rozhodují mezi formáty WAVE a FLAC (Knijff 2017).

## Formáty pro zpřístupňování zvukových dokumentů

Na formáty pro zpřístupnění nejsou kladeny tak vysoké nároky jako na formáty pro dlouhodobé uchovávání. Pro účely zpřístupnění se používají komprimované formáty, které několikanásobně zmenší velikosti původního souboru a umožní tak hladší a rychlejší přenos souborů po síti. Na volbu konkrétního formátu pro zpřístupnění má vliv efektivita komprese formátu a snadnost jeho použití pro uživatelskou komunitu (tj. podpora prohlížečů, podpora v aplikacích atd.). Vhodnými formáty jsou například MP3 a AAC, které patří do skupiny standardů MPEG.

<sup>23</sup> <http://www.lib.utah.edu/collections/digital/best-practices.php>

<sup>24</sup> <https://conservancy.umn.edu/pages/policies/>

<sup>25</sup> Průzkum 118 formátových politik severoamerických institucí z let 2012–2013 uvádí, že se v nich formát WAVE objevil jako preferovaný 80krát a formát AIFF 59krát.

<sup>26</sup> V danou chvíli se archivují ve formátu WAVE.

MP3, plným názvem MPEG-1 nebo MPEG-2 Audio Layer III, je dnes jedním z nejrozšířenějších komprimovaných formátů. Používá se pro přenos zvukových dokumentů přes internet i v přehrávačích hudby. Při kódování audia do MP3 se volí přenosová rychlost (bitrate), kterou bude signál kódován. Je tak možné ovlivnit výslednou velikost souboru i kvalitu zvuku. Čím vyšší přenosová rychlost, tím je vyšší výsledná velikost souboru a kvalita zvuku. Nejběžnější a de facto standardní je přenosová rychlost 128kbps pro monofonní zvuk a 256 kbps pro stereofonní zvuk, pro kvalitnější zvuk pak některé organizace volí vyšší bitrate. Například Stanford zpřístupňuje soubory MP3 s přenosovou rychlostí 160 kbps (mono) a 320 kbps (stereo) (Stanford University 201?).

AAC (Advanced Audio Coding) je novější standard MPEG, který nabízí větší rozsah vzorkovacích frekvencí a přenosových rychlostí a podporuje více kanálů než MP3. Má navíc efektivnější kompresi než formát MP3 – při stejné velikosti souboru poskytuje kvalitnější zvuk. Oproti formátu MP3 má ale také minimálně jednu nevýhodu. Oba formáty jsou chráněny patenty, avšak licencování softwaru pro soubory AAC je dražší a vlastníci patentu údajně i ostražitěji dohlížejí na dodržování licenční politiky (Fries c2005). Formát AAC pro zpřístupnění v současnosti využívá například Indiana university (Hardesty 2016). Od roku 2003 tento formát používá společnost Apple Computers pro online prodej hudby (Truesdell 2007, s. 516).

Na trhu existují další ztrátové formáty, které mají efektivnější kompresi a nejsou chráněny patenty (WMA, Ogg Vorbis apod.), avšak zatím se příliš nerozšířily. Fries (c2005, s. 210) uvádí, že se formát MP3 stal vítězem formátové války. Je totiž založen na otevřeném standardu, který byl přístupný vývojářům hardware a software dříve než dokumentace ostatních formátů. Přístroje tak tento formát začaly velmi záhy podporovat a uživatelé si na něj zvykli. Kvalita zvuku formátu MP3 je navíc dostačující a cena úložných kapacit klesá, není tedy zatím zájem o formáty poskytující stejnou kvalitu zvuku při menší velikosti souboru.

Pro Národní knihovnu ČR se zatím jako nejlepší volba jeví formáty WAVE a BWF pro dlouhodobé uchovávání a formát MP3 pro zpřístupnění. Jak bylo na základě výše uvedené analýzy zjištěno, jedná se o nejčastěji používané formáty v paměťových institucích. Další zvukové formáty, konkrétně například FLAC, AAC a Ogg Vorbis, však bude NK ČR nadále monitorovat a průběžně podrobovat posouzení. Jak správně poukazuje archivář audiovizuálních děl Dave Rice, jenž reaguje na přetrvávající preferenci formátů BWF/WAVE, aktuální osvědčené postupy („best practices“) nemusí vždy odpovídat aktuálnímu stavu technologií a jejímu nejlepšímu využití, a proto místo „osvědčené postupy“ užívá formulaci „prozatím dostatečné postupy“ („good-enough-for now practices“) (Rice 2013).

## Metadata

Pro účely zpřístupnění a dlouhodobého uchovávání je nutné digitální objekty doplnit metadatami, tj. informacemi popisujícími objekt, jeho původ, vlastnosti, strukturu, neporušenost, chování, související práva duševního vlastnictví a z nich vyplývající omezení přístupu apod. Metadata je možné rozdělit dle jejich účelu a obsahu do tří základních skupin: metadata popisná, administrativní a strukturální (Lazinger 2001). Zápis těchto metadat je vhodné provádět podle metadatových schémat (standardů), která určují elementy (metadatové prvky), jež je možné použít, a pravidla pro jejich zápis (Caplan 2003, s. 5). Použití těchto standardů usnadňuje výměnu dat mezi systémy, jejich konverzi a efektivní tvorbu metadat.

Popisná metadata popisují nebo identifikují informační zdroje a jsou důležitá zejména pro vyhledávání. Pro popisná metadata existuje několik metadatových schémat – univerzální schémata pro popis různých druhů dokumentů i schémata speciálně



vytvořená pro popis zvukových a audiovizuálních dokumentů. Je možné použít univerzální schéma Dublin Core<sup>27</sup>, které je jednoduché a přitom efektivní a umožňuje popsat široké spektrum objektů – od klasických tištěných dokumentů (knihy, časopisy atd.) po webové stránky, zvukové dokumenty i 3D objekty. Pro některé případy je však Dublin Core příliš stručný. Organizace IASA doporučuje schéma Dublin Core doplnit, nebo jej úplně nahradit, metadatovým schématem MODS<sup>28</sup> (International Association of Sound and Audiovisual Archives 2009, s. 24). Schéma MODS umožňuje podrobnější popis objektu než Dublin Core, jakož i více úrovní popisu (kolekce, strana desky, stopa), a stejně jako schéma Dublin Core může být použito k popisu různých typů dokumentů.

Pro popis zvukových dokumentů je možné dále použít metadatové schéma PBCore<sup>29</sup>. Toto schéma bylo vytvořeno speciálně pro popis zvukových a audiovizuálních dokumentů a mělo by tak lépe než dosud zmíněná schémata vyjádřit specifické charakteristiky těchto dokumentů.<sup>30</sup> Primárně bylo toto schéma určeno pro potřeby rozhlasových a televizních stanic. Dnes schéma PBCore spravuje organizace American Archive of Public Broadcasting, která jej od roku 2013 používá pro popis svých dokumentů. Schéma PBCore vychází ze schématu Dublin Core a je vyjádřené v jazyce XML. (Brylawski 2015, s. 84). V případě potřeby podrobnějšího popisu je do schématu PBCore možné vložit další schémata vyjádřená jazykem XML – například schémata PREMIS<sup>31</sup> a AES57<sup>32</sup>. Za přednost PBCore se považuje jeho jednoduchost a čitelnost pro člověka, jenž má s metadaty malé nebo žádné zkušenosti. Pro rozhlasové a televizní stanice je dostatečně komplexní. V paměťových institucích, zdá se, není toto schéma aktuálně příliš rozšířené.

Kombinace schémat Dublin Core a MODS je nyní součástí metadatové specifikace pro obrazová data v projektu Národní digitální knihovna (NDK) a bude pravděpodobně použita i pro ukládání zvukových dokumentů, tak, jak to doporučuje směrnice organizace IASA (International Association of Sound and Audiovisual Archives 2009). Jedná se patrně o nejčastěji používaná schémata pro popisná metadata v paměťových institucích, což usnadní jejich případnou výměnu mezi systémy a budoucí konverze do jiných formátů. Schéma MODS je dostatečně komplexní, a speciálně u schématu Dublin Core lze počítat s širokou podporou v různých softwarových systémech. Dalším důvodem pro toto rozhodnutí je katalogizační praxe. V případě, že digitalizované dokumenty jsou již popsány v knihovních systémech ve formátu MARC, je momentálně snazší je převést do schématu MODS než do schématu PBCore. Záznam v MODS je pak možné kdykoliv, pokud to dovolí časové možnosti, transformovat do schémat Dublin Core a PBCore.

Administrativní metadata slouží ke správě digitálních dat. Nesou informace o původu dat, o jejich digitálním životním cyklu, o právech, technických vlastnostech souborů apod. Administrativní metadata je dále možné rozdělit na metadata technická, archivační a metadata popisující práva duševního vlastnictví.

Technická metadata uchovávají informace o technických vlastnostech digitálních objektů, o softwarovém a hardwarovém prostředí, ve kterém vznikly či byly upraveny, apod. Tyto informace pak umožňují vyhledat dokumenty na základě jejich technických

<sup>27</sup> <http://dublincore.org/>

<sup>28</sup> <http://www.loc.gov/standards/mods/>

<sup>29</sup> <http://pbcore.org/>

<sup>30</sup> Jako příklad se uvádí mimo jiné popis zvukové nahrávky, která je uložena na více nosičích, nebo popis epizod seriálu apod. (ARSC, s. 85).

<sup>31</sup> <http://www.loc.gov/standards/premis/>

<sup>32</sup> <http://www.aes.org/publications/standards/search.cfm?docID=84>,  
<http://www.aes.org/standards/schemas/aes57-2011-08-27.xsd>

vlastností a v případě nesouladu těchto vlastností s aktuální praxí provést ochranná opatření k nápravě, například migraci zastaralého formátu do nového (Hutař 2012, s. 80–81).

Pro zápis technických metadat u zvukových dokumentů je možné použít několik metadatových standardů vytvořených speciálně pro zvukové dokumenty. Jedním z nich je metadatové schéma AudioMD, jež v průběhu let 2001–2003 vytvořila Kongresová knihovna. Původně bylo koncipováno jako schéma prozatímní. Až do roku 2011, kdy vyšlo schéma nové, bylo schéma AudioMD údajně nejlepším schématem pro popis technických vlastností (DeRidder 2015). Dnes toto schéma používají například Finská Národní digitální knihovna (Kansallinen digitaalinen kirjasto 2016) a knihovny Texaské univerzity (University of Texas Libraries, 2012; Cofield 2016). Knihovny Texaské univerzity se chystají na evaluaci schématu AES57 (Cofield 2016).

V roce 2011 vydala organizace Audio Engineering Society (AES) metadatový standard AES57, který by měl nahradit schéma AudioMD (Technical metadata c2016). Toto schéma umožňuje velmi podrobně popsat digitální objekt i jeho původní zdroj (z technických schémat zde představených umožňuje nejpodrobnější popis dokumentu). Schéma AES57 dnes používá například Harvardova univerzita a University of Alabama Libraries (DeRidder 2015).

Také již výše uvedené metadatové schéma PBCore dokáže zachytit technické vlastnosti digitálního audia. Oproti AES57 a AudioMD je jednodušší pro pochopení, ale méně podrobné a méně strukturované a bylo by tedy patrně nutné jej doplnit vložením schématu AES57.

Budoucnost schématu AudioMD se zdá být nejistá. Určitým indikátorem současného a budoucího postavení schématu AudioMD může být i jeho absence v příručce organizace ARSC, jež vznikla ve spolupráci s Kongresovou knihovnou<sup>33</sup>. Příručka pro technická metadata uvádí pouze schémata AES57 a PBCore. Schéma PBCore neumožňuje dostatečně podrobný technický popis. V současné situaci se tedy jeví jako nejlepší volba pro popis technických metadat schéma AES57, které je velmi obsáhlé, což odpovídá aktuální praxi v projektu Národní digitální knihovny, kdy je snaha zachytit maximální množství informací.

Technická metadata budou z části plněna pomocí extraktorů metadat JHOVE nebo FITS.

Archivační metadata obsahují informace nezbytné pro dlouhodobé uchování dokumentů. Jejich úkolem je podpořit životnost, použitelnost, autenticitu a srozumitelnost objektu i v budoucnosti (PREMIS Editorial Committee, 2015, s. 1). Jsou to informace dokumentující okolnosti vzniku digitálního dokumentu, informace o akcích, které byly provedeny s uloženými daty a metadaty, informace o interních a externích vztazích objektů apod. V projektu NDK se pro zápis archivačních metadat u obrazových dokumentů používá metadatové schéma PREMIS. Toto schéma bude použito i u dokumentů zvukových a bude rozšířeno o další nezbytné prvky, protože u zvukových dokumentů dnes neexistuje schéma, jež by umožňovalo zápis metadat souvisejících se vznikem digitálního zvukového dokumentu<sup>34</sup>. Jedná se o informace o použitých zařízeních (skener, přehrávač, předzesilovač...) a jejich nastaveních.<sup>35</sup>

<sup>33</sup> ARSC Guide to Audio Preservation, <https://www.clir.org/pubs/reports/pub164/pub164.pdf>

<sup>34</sup> U obrazových dokumentů bylo možné tyto údaje zapsat do schématu MIX, jenž slouží pro zápis technických metadat.

<sup>35</sup> Tyto informace mohou být důležité například v případě, kdy se zjistí špatná funkčnost některého z použitých zařízení (skeneru, kodeku). Umožňují vyhledat dokumenty, při jejichž tvorbě bylo toto zařízení použito.

Digitální objekty jsou často tvořeny více soubory stejného i různého typu – několik verzí zvukového souboru, metadata, obrazová data (např. obálka, deska, booklet) apod. Strukturální metadata vyjadřují, jak spolu jednotlivé objekty souvisejí, a umožňují tak zobrazení dokumentu ve smysluplné podobě. Strukturu dokumentu je možné vyjádřit, a pravděpodobně tomu tak u zvukových dokumentů bude, pomocí schématu METS a jeho strukturní mapy. Schéma METS bude také kvůli přehlednosti a jednodušší práci použito jako kontejner pro vybraná metadatová schémata.

Vedle zvukových dokumentů a jejich metadat budou součástí dat přijímaných k dlouhodobému uchovávání i obrazové dokumenty doprovázející zvukové dokumenty, například obal, brožura (booklet), fotografie fyzického hudebního dokumentu (desky apod.). Tato obrazová data budou vytvářena podle aktuálních specifikací pro obrazová data v projektu NDK. Obrazy budou dodávány ve formátu JPEG2000, a to pravděpodobně ve dvou verzích (v archivační a zpřístupňující kvalitě). Pro popis jejich technických vlastností bude použit standard MIX.

## Validace souborových formátů a extrakce technických metadat

Na rozhodování o archivačním souborovém formátu má vliv i existence nástrojů, které jej dokáží identifikovat, validovat a charakterizovat. Testováno bylo několik nástrojů. Pro proces identifikace byly testovány nástroje DROID<sup>36</sup>, Siegfried<sup>37</sup> a FIDO<sup>38</sup>. Všechny tři nástroje využívají pro identifikaci registr PRONOM<sup>39</sup>. Nástroj FIDO nerozpoznal jednu z podverzí formátu WAV, resp. ji označil za verzi jinou, proto jej pro použití v našich procesech zatím dále nezvažujeme. Nástroje DROID a Siegfried se ve všech případech shodovaly, k identifikaci bude tedy zvolen pravděpodobně jeden z nich<sup>40</sup>.

Pro validaci je možné použít nástroj JHOVE<sup>41</sup>, jehož modul WAVE-hul rozezná formáty BWF a WAVE a jejich jednotlivé verze (WAVEFORMATEX, PCMWAVEFORMATEX apod.).

Pro extrakci metadat existuje opět několik nástrojů. Testovány byly nástroje JHOVE, MediaInfo<sup>42</sup>, FITS<sup>43</sup>, Apache Tika<sup>44</sup>, Exiftool<sup>45</sup> a poměrně nový nástroj PET<sup>46</sup>. Nejlépe si vedl nástroj JHOVE, který je ve výpisu vlastností velmi detailní, přesný<sup>47</sup> a oproti ostat-

<sup>36</sup> DROID, Digital Record Object Identification, <http://www.nationalarchives.gov.uk/information-management/manage-information-policy-process/digital-continuity/file-profiling-tool-droid/>

<sup>37</sup> Siegfried, <http://www.itforarchivists.com/siegfried>

<sup>38</sup> FIDO, Format Identification for Digital Objects, [http://coptr.digipres.org/FIDO\\_\(Format\\_Identifier\\_for\\_Digital\\_Objects\)](http://coptr.digipres.org/FIDO_(Format_Identifier_for_Digital_Objects))

<sup>39</sup> Národní archiv Velké Británie spravuje formátovou knihovnu PRONOM, která obsahuje informace o jednotlivých souborových formátech včetně tzv. signature files, které slouží k identifikaci formátů prostřednictvím identifikačních nástrojů, jako jsou DROID, FIDO a Siegfried.

<sup>40</sup> Pokud další testy neukáží jinak, bude preferovaným nástrojem pro identifikaci nástroj DROID, protože ho spravuje uznávaná organizace v oblasti dlouhodobého uchovávání (The National Archives).

<sup>41</sup> JHOVE, JSTOR/Harvard Object Validation Environment, <http://jhove.openpreservation.org/>

<sup>42</sup> MediaInfo, <https://mediaarea.net/cs/MediaInfo>

<sup>43</sup> FITS, File Information Tool Set, <http://projects.iq.harvard.edu/fits/home>

<sup>44</sup> Apache Tika, <https://tika.apache.org/>

<sup>45</sup> Exiftool, <http://www.sno.phy.queensu.ca/~phil/exiftool/>

<sup>46</sup> PET, PERICLES Extraction Tool, [http://coptr.digipres.org/PERICLES\\_Extraction\\_Tool\\_\(PET\)](http://coptr.digipres.org/PERICLES_Extraction_Tool_(PET))

<sup>47</sup> Nástroj MediaInfo v našem testu zaokrouhloval čas nahrávky (duration).

ním nástrojům podává spolu s extrahovanými metadaty i zprávu o výsledku validace. Je možné též použít nástroj FITS, který obsahuje nástroje DROID, JHOVE a MediaInfo. Bylo zjištěno, že například v University of Alabama používají pro validaci, charakterizaci a identifikaci právě nástroj FITS. Na jeho výstupy namapovali schéma AES57 a mohou tak automatizovaně plnit technická metadata a zároveň provádět identifikaci a validaci formátu (DeRidder 2015).

## **Závěr**

Zvukové dokumenty jsou důležitou součástí historie každé země, stejně jako tištěné materiály dokumentují její vývoj. Globálně se jejich ochraně a uchovávání věnuje pozornost již více než 10 let, zahraniční zkušenosti a postupy tak mohou být bohatým zdrojem informací pro práci s těmito dokumenty u nás. V ČR nyní dochází k jejich výběrové inventarizaci a v brzké době se počítá i s jejich digitalizací. Výhledově tedy bude naše zvukové kulturní bohatství do určité míry zpřístupněno uživatelům a uchováno pro budoucí generace, k čemuž je nezbytná právě příprava archivačního standardu, stejně jako formátů pro tuto dlouhodobou ochranu. V představené studii byly shrnuty informace o využívaných standardech pro archivaci i zpřístupnění (jak pro datové formáty, tak pro metadata) a navrženy vhodné postupy pro přípravu národního standardu na základě postupů užívaných v globální komunitě.

## Použitá literatura

- BENTLEY HISTORICAL LIBRARY. Format Conversion Strategies for Long-Term Preservation. *Bentley historical library University of Michigan* [online]. c2016 [cit. 2016-12-05]. Dostupné z: <http://bentley.umich.edu/about/what-we-do/digital-curation-strategies-and-procedures/680-2/>.
- BINGHAMTON UNIVERSITY. UNIVERSITY LIBRARIES. Digital preservation: file format support [online]. Binghamton, NY: Binghamton University Libraries, c2016. Dostupné z: <http://www.binghamton.edu/libraries/technology/digital-preservation/levels-based-on-format.html>.
- BRYLAWSKI, SAM etc., 2015. *ARSC Guide to Audio Preservation*. 240 s. ISBN 978-1-932326-50-5. Dostupné též z: <https://www.clir.org/pubs/reports/pub164/pub164.pdf>.
- CAPLAN, Priscilla, 2003. *Metadata fundamentals for all librarians*. Chicago: American Library Association, s. 5. ISBN 08-389-0847-0.
- CDP DIGITAL AUDIO WORKING GROUP. *Digital audio best practices*. Version 2.1. 2006. Dostupné také z: <http://ucblibraries.colorado.edu/systems/digitalinitiatives/docs/digital-audio-bp.pdf>.
- COFIELD, Melanie C. RE: AudioMD for technical metadata [elektronická pošta]. Message to natalie.ostrakovakp.cz. 26. července 2016 21:11 [cit. 2016-07-28]. Osobní komunikace.
- COLORADO STATE UNIVERSITY. LIBRARIES. CSU Digital repository preservation & format support policy [online]. 2015. Dostupné z: <http://lib.colostate.edu/repository/csu-digital-repository-preservation-format-support-policy>.
- CONSORTIUM OF ACADEMIC AND RESEARCH LIBRARIES IN ILLINOIS. Guidelines for the creation of digital collections: digitization best practices for audio [online]. 2014. Dostupné z: [https://www.carli.illinois.edu/sites/files/digital\\_collections/documentation/guidelines\\_for\\_audio.pdf](https://www.carli.illinois.edu/sites/files/digital_collections/documentation/guidelines_for_audio.pdf).
- CROWE, Jason. *Audio File Format for Preservation* [elektronická pošta]. Message to: natalie.ostrakovakp.cz. 8. srpna 2016 1:11 [cit. 2016-11-28]. Osobní komunikace.
- ČESKO. MINISTERSTVO KULTURY. Rozhodnutí ministra kultury č. 3/2015: ze dne 15. června 2015 [online]. 2015. Praha: Ministerstvo kultury ČR [cit. 2016-11-11]. Dostupné také z: <http://www.mzk.cz/sites/mzk.cz/files/souboryMZK/pdf/doc20150625185154.pdf>.
- ČESKO. MINISTERSTVO KULTURY. *Plán implementace Státní kulturní politiky na léta 2015–2020* [online]. 2015. Praha: Ministerstvo kultury ČR. Dostupné z: <https://www.mkcr.cz/statni-kulturnipolitika-69.html>.
- DERIDDER, Jody L. Generating Standardized Audio Technical Metadata: AES57. *Code4lib* [online]. 2015, **30** [cit. 2016-12-05]. ISSN 1940-5758. Dostupné z: <http://journal.code4lib.org/articles/10828>.
- FLORIDA CENTER FOR LIBRARY AUTOMATION. *Florida Voices* [online]. c2008 [cit. 2016-06-10]. II. Preservation. Dostupný z: <http://www.fcla.edu/FloridaVoices/preservation.shtm#Recommended>.
- FRIES, Bruce a Marty FRIES, c2005. *Digital audio essentials*. Sebastopol, CA: O'Reilly. 376 s. ISBN 05-960-0856-2.
- HARDESTY, Julie. L. Re: questions about sound metadata [elektronická pošta]. Message to: filip\_sirm.cz. 19. května 2016 17:03 [cit. 2016-11-28]. Osobní komunikace.
- HOROVÁ, Iva, Helena NOVOTNÁ a Filip ŠÍR, 2016. *Za málo peněz hodně muziky aneb co se děje v oblasti zvukových dokumentů u nás i v zahraničí*. In: INFORUM 2016: 22. ročník konference o profesionálních informačních zdrojích, Praha 24.–25. května 2016 [online]. Praha: Albertina icome Praha. [cit. 2016-11-11]. ISSN 1801-2213. Dostupné z: <http://www.inforum.cz/pdf/2016/sir-filip.pdf>.
- HUTAŘ, Jan, 2012. *Digitalizace, popis pomocí metadat a jejich formáty*. Praha, 244 s. Disertační práce. Univerzita Karlova v Praze, Ústav informačních studií a knihovnictví.
- INTERNATIONAL ASSOCIATION OF SOUND AND AUDIOVISUAL ARCHIVES. *Guidelines on the Production and Preservation of Digital Audio Objects*, ed. by Kevin Bradley. Second edition 2009. ISBN 978-91-976192-3-3. Dostupné též z: <http://www.iasa-web.org/audio-preservation-tc04>.
- J. WILLARD MARRIOTT LIBRARY. Digitization Best Practices. *J. Willard Marriott Library – The University of Utah* [online]. Salt Lake City: The University of Utah. Poslední aktualizace 2015-06-25 [cit. 2016-08-08]. Dostupné z: <http://www.lib.utah.edu/collections/digital/best-practices.php>.

KANSALLINEN DIGITAALINEN KIRJASTO. File Formats [online]. Finsko, 2016. Version 1.4.0. Dostupné z: <http://www.kdk.fi/images/tiedostot/NDL-File-Formats-v1.4-en.pdf>.

KNIJFF, Johan van der. Breaking WAVes (and some FLACs), 2017. In: *Shared solutions for effective and efficient digital preservation – Open Preservation Foundation* [online]. The Open Preservation Foundation. [cit. 2017-03-27]. Dostupné z: <http://openpreservation.org/blog/2017/01/04/breaking-waves-and-some-flacs/>.

Koncepce Moravské zemské knihovny na období 2015–2025 [online]. 2015. Brno: Moravská zemská knihovna v Brně [cit. 2016-11-11]. Dostupné z: <https://www.mzk.cz/oknihovne/dokumenty/koncepce-moravske-zemske-knihovny-na-obdobi-2015-2025>.

LAZINGER, Susan S., 2001. Digital preservation and metadata: history, theory, practice. Englewood: Libraries Unlimited. s. 142–143. ISBN 1-56308-777-4.

LIBRARY AND ARCHIVES CANADA. Guidelines on File Formats for Transferring Information Resources of Enduring Value. *Library and Archives Canada* [online]. 2014. Poslední aktualizace 2015-02-06 [cit. 2016-08-08]. Dostupné z: <http://www.bac-lac.gc.ca/eng/services/government-information-resources/guidelines/Pages/guidelines-file-formats-transferring-information-resources-enduring-value.aspx#p>.

LIBRARY OF CONGRESS. AIFF (Audio Interchange File Format). *Sustainability of Digital Formats Planning for Library of Congress Collections* [online]. 2014 [cit. 2016-12-05]. Dostupné z: <http://www.digitalpreservation.gov/formats/fdd/fdd000005.shtml>.

LIBRARY OF CONGRESS. *Library of Congress Recommended Formats Statement 2016-2017* [online]. 2016. Dostupné z: <http://www.loc.gov/preservation/resources/rfs/RFS%202016-2017.pdf>.

LIBRARY OF CONGRESS. *Sustainability of Digital Formats: Planning for Library of Congress Collections* [online]. Aktualizace 2013 [cit. 2016-06-10]. Sustainability Factors. Dostupné z: <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.

THE NATIONAL ARCHIVES AND RECORDS ADMINISTRATION. NARA 2014-04: Appendix A, Revised Format Guidance for the Transfer of Permanent Electronic Records – Tables of File Formats. In: *National Archives and Records Administration* [online]. The U.S. National Archives and Records Administration, 2014 [cit. 2016-07-11]. Dostupné z: <https://www.archives.gov/records-mgmt/policy/transfer-guidance-tables.html>.

PREMIS EDITORIAL COMMITTEE. 2015. PREMIS Data Dictionary for Preservation Metadata [online]. Version 3.0. Washington: PREMIS Editorial Committee, June 2015 [cit. 2016-08-08]. Dostupné z: <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>.

Remo Software. WAV File. Remo Software Glossary. [Online] 2015. Dostupné z: <http://www.remosoftware.com/glossary/wav-file>.

RICE, Dave. FLAC in the archives. In: *Dave rice | archivist | technologist* [online]. 2013 [cit. 2016-06-05]. Dostupné z: <http://dericed.com/2013/flac-in-the-archives/>.

RIMKUS, Kyle et al. *Digital Preservation File Format Policies of ARL Member Libraries: an analysis*. D-Lib Magazine. 20 (3/4) 2014. doi:10.1045/march2014-rimkus. Dostupné z: <http://www.dlib.org/dlib/march14/rimkus/03rimkus.html>.

ROG, Judith and Caroline van WIJK. 2008. *Evaluating File Formats for Long-term Preservation* [online]. 2008 [cit. 2011-05-02]. Dostupné z: [http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf).

STANFORD UNIVERSITY. Capture specs. *Stanford University Libraries* [online]. Stanford, [2017] [cit. 2016-06-27]. Dostupné z: <https://library.stanford.edu/research/digitization-services/labs/stanford-media-preservation-lab/capture-specs>.

STATE LIBRARY NEW SOUTH WALES. Sound recordings: *State library New South Wales* [online]. Sydney: State Library of NSW, June 25th 2015 [cit. 2016-08-08]. Dostupné z: <http://www.sl.nsw.gov.au/public-library-services/sound-recordings>.

ŠÍR, Filip a Petr ŽABIČKA, 2013. *Metodika pro digitalizaci a on-line zpřístupňování gramofonových nahrávek pro pamětové instituce*. Brno. Dostupné také z: <https://docs.google.com/viewer?a=v&pid=site-s&srcid=ZGVmYXVsdGRvbWpbnxuYXJvZG5pZm9ub3Rla2F8Z3g6NWJmNDFlMjgWmJmNDkZA>.

TASMANIAN ARCHIVE + HERITAGE OFFICE. *Digital Preservation Formats: State Records Guideline No 19*. 2015. Dostupné také z: <https://www.informationstrategy.tas.gov.au/Records-Management-Principles/Document%20Library%20%20Tools/Guideline%2019%20Digital%20Preservation%20formats.pdf>.

Technical metadata. *JISC Digital Media* [online]. c2016 [cit. 2016-12-05]. Dostupné z: <http://www.webarchive.org.uk/wayback/archive/20160101152122/http://www.jiscdigitalmedia.ac.uk/infokit/audiovisual-digitisation/technical-metadata>.

TRUEDELL, Cliff. 2007 *Mastering digital audio production the professional music workflow with Mac OS X*. Hoboken, N. J: Wiley, s. 516. ISBN 978-047-0165-768.

UNESCO/UBC (2012). *VANCOUVER DECLARATION: The Memory of the World in the Digital Age: Digitization and Preservation*. UNESCO, Vancouver. Dostupné z: [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco\\_abc\\_vancouver\\_declaration\\_en.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_abc_vancouver_declaration_en.pdf).

UK Data Archive. FILE FORMATS & SOFTWARE. *UK Data archive* [online]. c2002-2016 [cit. 2016-12-05]. Dostupné z: <http://data-archive.ac.uk/create-manage/format/formats>.

UNIVERSITY OF EDINBURGH. *Edinburgh DataShare: Recommended File Formats* [online]. Poslední aktualizace April 2015 [cit 2016-08-08]. Dostupné z: [http://www.ed.ac.uk/files/atoms/files/recommended\\_file\\_formats-apr2015.pdf](http://www.ed.ac.uk/files/atoms/files/recommended_file_formats-apr2015.pdf).

UNIVERSITY OF MINNESOTA. LIBRARIES. DIGITAL CONSERVANCY. *Policies and Guidelines* [online] c2014. Dostupné z: <https://conservancy.umn.edu/pages/policies/#preservation>.

UNIVERSITY OF TEXAS LIBRARIES. *University of Texas Libraries Human Rights Documentation Initiative Metadata Guidelines for Audio*. Version 1. 1. 2012. Dostupné také z: [https://www.lib.utexas.edu/schema/Audio\\_Metadata\\_Guidelines\\_v1.pdf](https://www.lib.utexas.edu/schema/Audio_Metadata_Guidelines_v1.pdf).

WHIBLEY, SIMON. *Wav Format Preservation Assessment*. 2016. British Library. Version 1.0. Dostupné z: [http://wiki.dpconline.org/images/4/46/WAV\\_Assessment\\_v1.0.pdf](http://wiki.dpconline.org/images/4/46/WAV_Assessment_v1.0.pdf).

XIPH.ORG FOUNDATION. *Flac free lossless audio codec* [online]. Aktualizace 27. 11. 2014 [cit. 2016-06-15]. Dostupný z: <https://xiph.org/flac/>.

YORK UNIVERSITY LIBRARIES. Digitization best practices. *York University* [online]. Toronto: York University, 2015. [cit. 2016-08-08]. Dostupné z: <http://www.library.yorku.ca/web/collections/digitalscholarship/digitization-best-practices/>.

OSTRÁKOVÁ, Natalie a Filip ŠÍR. Zvukové dokumenty ve fondech paměťových institucí v kontextu dlouhodobého uchovávání v ČR. Přípravná studie NK ČR k možnosti dlouhodobého uložení digitalizovaných dat. *Knihovna: knihovnická revue*. 2017, **28**(1), 5–19. ISSN 1801-3252.