

Správa a archivace audiovizuálního materiálu

Bohuš Získal

Ústav informačních studií a knihovnictví, Filosofická fakulta University Karlovy
bohus.ziskal@ff.cuni.cz

INFORUM 2009: 15. konference o profesionálních informačních zdrojích
Praha, 27. – 29. 5. 2008

Abstrakt:

S rostoucí dostupností technologií roste i množství audiovizuálního materiálu použitelného jako informační zdroj nebo přímo pořízeného za účelem ukládání informací, například pro dokumentování výstupů vědeckého experimentu. Ačkoli v současné době jsou technologie pro elektronické vytváření, zpracování, sdílení a distribuci audiovizuálního obsahu široce používány, řešení pro jeho ukládání a uchovávání odrážejí zejména potřeby jeho producentů. Tento příspěvek popisuje problematiku práce s audiovizuálním materiálem jako s informačním zdrojem, zaměřuje se na obecně použitelné způsoby jeho správy a dlouhodobé archivace. Poskytuje přehled a výhody aktuálně dostupných metod vytváření a správy doprovodných informací v textové podobě (metadat). Na definici standardu MXF a jeho variantách implementovaných v komerčně dostupných zařízeních je demonstrován přínos a omezení jednotného souborového formátu pro ukládání obsahu spolu s metadaty. V příspěvku jsou uvedeny metody extrahování textových informací z audiovizuálního materiálu a možnosti formátu MPEG-4 pro ukládání metadat identifikujících obsah. Závěrečná část je věnována popisu systému pro správu digitálního obsahu (DAM – Digital Asset Management) určeného pro audiovizuální materiál. Na dostupných technologiích a trendech jsou demonstrovány postupy při jeho návrhu a provozování, které mají klíčový význam pro zpracování a archivaci informací a jsou vhodné i pro instituce, které se primárně vytvářením a zpracováním audiovizuálního obsahu nezabývají.

Specifická povaha audiovizuálního materiálu

Audiovizuální materiál jako datový objekt je tvořený digitální reprezentací vizuální a zvukové složky, k jejichž smyslově vnímatelné prezentaci je nutné technické zařízení schopné převést data na příslušnou fyzikální reprezentaci. Zdrojem těchto dat může být jak zařízení pro převod a kódování fyzikálních veličin (světla nebo zvuku) do digitální podoby (kamera, mikrofon), tak i výpočetní systém (počítačem generovaná grafika nebo zvuk).

Aby bylo možné data v digitální podobě dekodovat, musí být známá jejich reprezentace a uspořádání, v technické terminologii jejich formát. Samotný termín formát (angl. Format) má řadu významů, které se někdy specifikují podle použití nebo způsobu zacházení. Formát může označovat jak uspořádání dat, konkrétní reprezentaci - typ souboru (formát AVI) , tak i fyzickou podobu nosiče (formát DVD), nebo kombinaci nosič plus médium (formát XDCAM). V audiovizuální oblasti se dále používají termíny akviziční formát (acquisition format) pro označení formátu (nosiče + dat) určeného k pořízení obrazového nebo zvukového záznamu, a distribuční formát, záznamový nebo studiový formát, což jsou formáty určené pro konkrétní použití.

Informace obsažené v audiovizuálním materiálu mají specifickou povahu, obraz a zvuk má prezentační charakter na rozdíl od informačního charakteru textu (Meyrowitz, 2006) a velkou roli zde hraje kontext. Doprovodné informace tak mohou být zásadní pro interpretaci informace obsažené v obraze nebo zvuku, například čas a způsob pořízení záznamu, lokalita nebo jména zobrazených osob. Některé z těchto informací nelze získat sebepečlivější analýzou dat. Rozlišování mezi

informacemi, které se dají získat analýzou obrazové a zvukové složky záznamu, a doprovodnými kontextovými informacemi, je klíčové pro práci audiovizuálním materiálem jako informačním zdrojem. Například převod mluveného slova ve zvukové stopě na text a následné vyhledávání lze provést v zásadě s libovolným časovým odstupem po pořízení záznamu (pokud je záznam dekódovatelný), přiřazení jmen osob k zobrazeným tvářím je možné jen za podmínky existence reference. Pro doprovodné informace se často užívá termín metadata, pro materiál samotný potom termín esence. Metadata a esence pak dohromady tvoří celkový obsah, dále tedy budu termíny používat v tomto významu. Pro metadata, která lze získat analýzou esence budu dále používat termín **odvozená**, zbylá metadata budu nazývat **primární**.

Metadata lze klasifikovat do různých kategorií podle obsahu nebo jejich vztahu k materiálu samotnému. Pro příklad zde uvádím jedno možné členění metadat:

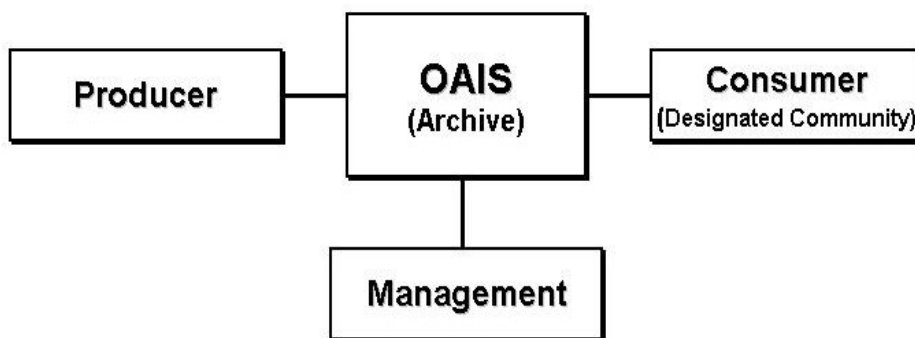
- Technická – většinou informace potřebné pro zpracování a rekonstrukci esence, někdy ale mohou poskytovat technický kontext k esenci (například typ objektivu při snímání)
- Popisná – informace vážící se k samotné esenci, mohou být jak primární, tak odvozená
- Administrativní – váží se k použití esence, například autorská práva

Způsob pořizování, správy a archivace a audiovizuálního materiálu by z tohoto pohledu měl preferovat zachování všech typů metadat, která nelze získat z esence.

Správa a archivace audiovizuálního materiálu ve vztahu k jeho vytváření

Z hlediska metadat je důležitý vztah producenta obsahu k procesu archivace.

V konceptuálním modelu archivu OAIS jsou definovány role poskytovatele obsahu, správce obsahu a cílové skupiny uživatelů obsahu (Designated Community). Tyto role lze mohou být přiřazeny v rámci jedné organizace, jak je tomu může být například v případě televizní stanice, nebo mohou být externí vůči samotnému archivu tak, jak to popisuje model OAIS (Sawyer et al.,2002).



OAIS model prostředí (začlenění archivu do prostředí)

Z hlediska přístupu k doprovodným informacím (metadatům), které nelze analýzou esence získat, můžeme na typ vztah archivu k produkci obsahu pohlížet jako na:

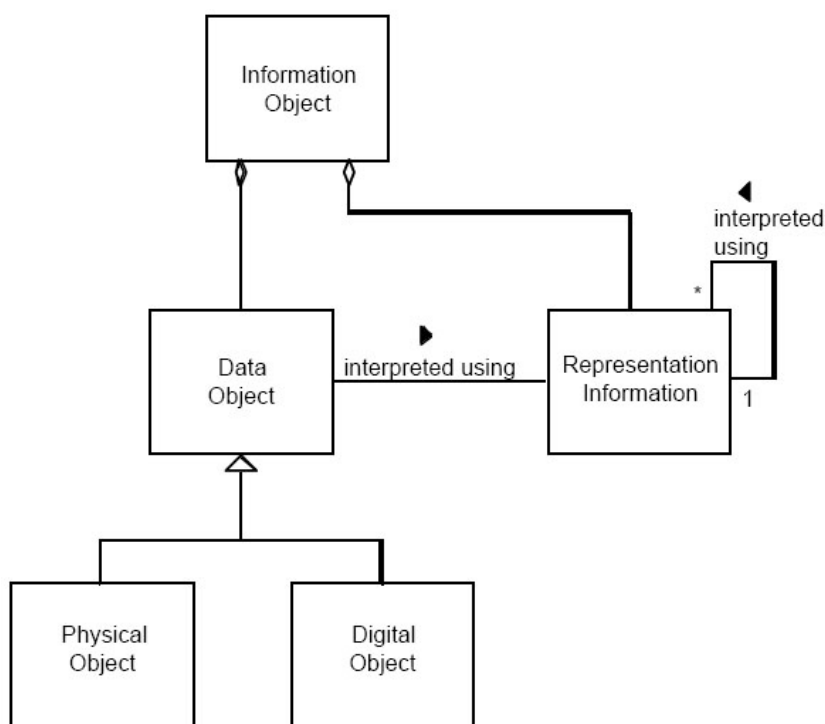
- **Produkční** – archiv je přímo navázaný na produkci obsahu
- **Sběrný** – data přicházejí do archivu z různých distribučních kanálů

Ve smyslu tohoto rozdělení producenti audiovizuálního obsahu budují a provozují většinou

produkční archivy s přímým přístupem k primárním metadatům, zatímco tradiční archivní instituce se zaměřují na sběrné archivy. U produkčních archivů pak splývají role producenta, managementu a uživatelů, jelikož externí uživatelé většinou nemají přímý přístup k archivu (obsah se k nim dostává prostřednictvím distribučních kanálů). Pokud provozovatel sběrného archivu nezajistí přístup k primárním metadatům, může být informační hodnota archivovaného obsahu omezená bez ohledu na kvalitu archivace nebo práce s odvozenými metadaty. V praxi se často setkáváme se situací, kdy řada primárních metadat není archivována spolu s esencí, ale je uložena v systému DAM (Digital Asset Management). Záleží potom na daném producentovi, jakým způsobem zajišťuje svazování těchto metadat při exportu z archivu. Moderní formáty, jako je dále zmíněný MXF, dovolují přenášet esenci i metadata v jednom souboru.

Obecné standardy pro ukládání a distribuci metadat

Strukturu metadat a jejich koncepční členění ovlivňuje řada faktorů. V první řadě je to samotný koncept informačního modelu, který vyplývá z pohledu na samotný účel archivace. Zjednodušeně lze odlišná východiska ilustrovat například na zájmu producenta obsahu, který archiv používá k redistribuci a recyklaci materiálů, a na cílech instituce zaměřené na specifické užití obsahu, například pro dokumentaci vědeckých pokusů. Snaha a standardizaci těchto konceptů potom leží zejména na straně profesních sdružení, které mají dlouhodobé uchovávání informací v základní náplni. Dva nejrozšířenější koncepční modely vychází z CIDOC-CRM (Conceptual Reference Model od International Committee for Documentation) a FRBR (Functional Requirements for Bibliographical Records), z hlediska systematického pohledu na dlouhodobou archivaci je potom standardizován ještě Reference Model for an Open Archival Information System (Sawyer et al.,2002).



Základní model informačního objektu podle OAIS

Z těchto modelů vycházejí jak ontologie ve smyslu „specifikace konceptualizace“ (Gruber, 1995), tak struktura informačních objektů, případně jejich taxonomie. Zmíněný OAIS model například termín metadata nepoužívá, protože nevystihuje přesně funkci žádného z jeho datových objektů (Sawyer et al.,2002). Koncept informačního modelu také koresponduje s architekturou archivu (např. Kahn/Wilensky 2006), která nemusí mít v případě empiricky vyvíjených řešení explicitně definovaný koncept. Na jedné straně tak stojí pojetí metadat jako struktury vycházející z určitého přístupu

k budování archivu, na druhé straně mohou být metadata vytvářena čistě účelově pro konkrétní potřeby určité aplikace. Vzhledem k tomu, že audiovizuální archivy budované pro potřeby producentů obsahu v masmediálním průmyslu vznikají postupně a jsou obvykle úzce navázané na výrobu, naplňují metadata zejména funkční potřebu bez jasně definovaného informačního modelu a v řadě případů jejich struktura neodpovídá žádnému z obecných standardů.

Standardy popisných metadat

V audiovizuálním průmyslu existuje řada proprietárních řešení, která většinou bývají navázána na konkrétní databázové (nejčastěji SQL nebo Oracle) či souborové (XML nebo HTML) řešení. Vzájemné mapování atributů mezi databázemi ve většině případů nepředstavuje technický problém, obtížnější je export relací a funkčních skupin. V oblasti audiovizuálního průmyslu jsou standardy navázány na ukládání a distribuci esence, což platí částečně pro MPEG-7, ale zejména pro AAF a dále popisovaný MXF. Kromě těchto formátů existuje řada standardů popisujících schémata uspořádání jednotlivých atributů pro konkrétní reprezentaci, například RDF, DCD nebo SOX (Object-oriented XML). Srovnáním jednotlivých standardů se zabývá řada studií (např. De Sutter et al., 2006).

Samostatnou oblast tvoří čistě distribuční schémata, která nedefinují konkrétní souborový formát, ale sadu parametrů pro jednotlivé skupiny metadat (příkladem je např. P/META). Pro výše zmíněné standardy je charakteristická orientace zejména na praktické produkční atributy, například u AAF je to schopnost přenášet popis rozpracovaného projektu včetně údajů o vrstvách esence a aplikovaných operacích. Pro samotné kódování metadat stojí za zmínku například standard METS (Metadata Encoding & Transmission Standard), zaměřený na XML formát, který respektuje dělení metadat na popisná, technická, administrativní a strukturální.

Pro potřeby distribuce je navrženo a v rámci EBU doporučováno schéma P/Meta, které je primárně určené pro výměnu esence spolu s metadaty a využívá formátu XML. Jelikož standard P/Meta sám o sobě nedefinuje konkrétní uspořádání databázové struktury ani soubory metadat, lze v rámci něj využít aktuální soubory atributů používané konkrétním producentem obsahu, pokud jsou v souladu s mezinárodními zvyklostmi (například s číselníky pro identifikaci žánrů používané v rámci EBU). P/Meta standard je jazykově a systémově nezávislý, proto se hodí i pro český jazyk.

MXF – Material eXchange Format

V audiovizuálním průmyslu proběhlo několik pokusů o standardizaci jednotného formátu. Specifikace v současné době nepropracovanějšího standardizovaného souborového formátu, který dostal název MXF, vznikala poměrně dlouho. Výsledek pokrývá řadu aplikací, o čemž svědčí i řada norem, které se k němu vztahují (SMPTE 377M – hlavní specifikace, implementační standardy jako SMPTE 390M – OP Atom, generické kontejnery – např. SMPTE 379M, popisy deskriptivních metadat - SMPTE 380M, atd.). Popis struktury MXF je velmi komplikovaný a zdaleka ne všechny specifikované části se v praxi používají. Zjednodušeně si lze formát MXF představit jako souborový kontejner (wrapper), který v sobě může nést primární video stopu, audio stopy, video v náhledovém rozlišení, statické obrázky (Thumbnails), popisné informace (metadata), časový kód a doprovodná data jako titulky nebo teletext. Jednotlivé datové struktury mohou být navázána na časový kód, například statický obrázek nebo titulek lze připojit k definovanému místu u videa. Není asi třeba zdůrazňovat praktické výhody takto pokročilého formátu, proto z něj i aktuálně nejrozšířenější záznamové formáty P2 (Panasonic) a XDCAM (Sony) vycházejí a nabízejí jak metadata, tak i video v náhledovém rozlišení či statické obrázky. Pro uživatele zůstává existence souborů MXF často skrytá, jelikož pokud například server pracuje interně ve formátu MXF, uživatel má v příslušné aplikaci k dispozici již rozbalenou informaci v podobě příslušných video a audio stop nebo například titulků (Wells et al., 2006). Vzhledem k tomu, že formát MXF dovoluje navazovat metadata i na specifické části obsahu (časový kód), je jejich struktura poměrně složitá. V zásadě je lze rozdělit na metadata v hlavičce souboru (Header metadata), metadata popisující vztah jednotlivých stop (Inter-track metadata) a metadat v rámci každé stopy (Intra-track metadata).

Způsoby využití a vytváření odvozených metadat

Z hlediska dlouhodobé archivace mají odvozená metadata jen malý význam, jelikož jejich případná ztráta nemá vliv na uchovávanou informaci. Pro přístup k esenci, prohledávání a další zpracování informací však mohou mít doprovodná metadata zásadní význam. Hlavním důvodem je, že prohledávání audiovizuálního obsahu je vzhledem k jeho povaze časově velmi nákladný proces. Řadu operací je mnohem snazší a efektivnější provádět s popisnými metadaty, která na příslušný audiovizuální obsah odkazují. U některých formátů, jako je výše zmíněný MXF, lze metadata navazovat přímo na časový údaj určující pozici v audiovizuální esenci. Pro prohledávání a manipulaci je takto navázaná informace velmi užitečná, je ovšem otázkou, zda obsah archivovat i takto navázanými odvozenými metadaty.

Analýza zvuku

Analýza zvuku může probíhat ve dvou rovinách, jednak jako převod mluveného slova do textu, nebo jako vyhledávání určitého zvuku nebo souboru zvuků podle vzoru (například zatroubení auta). Algoritmy pro převod na text se již v praxi používají zejména pro anglosaské jazyky a již uspokojivě fungují i pro český jazyk. Jelikož je analýza stále citlivá na konkrétního mluvčího, používá se trik s opakováním v reálném čase, kdy mluvčí, jehož přednes vykazuje nízkou chybovost (systém se na něj naučí), opakuje v reálném čase mluvené slovo ze zvukové stopy.

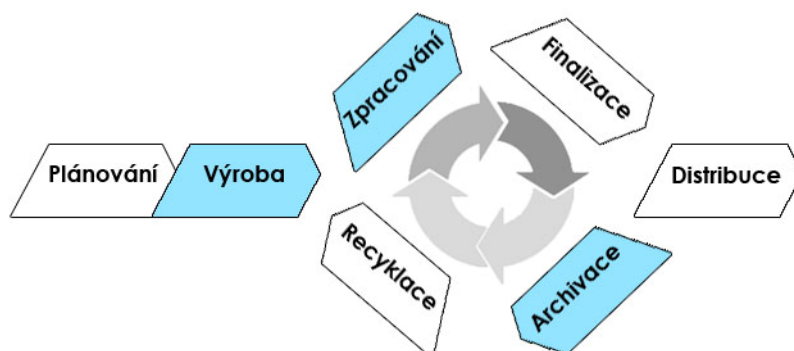
Analýza obrazu

Stejně jako analýza zvuku, je i analýza obrazu již součástí existujících komerčních řešení. Používá se zejména vyhledávání shody s určitým vzorem (pattern recognition). Typická aplikace je vyhledávání tváří podle fotografické předlohy. Přesnost samozřejmě závisí na kvalitě obrazu, přesto moderní algoritmy vykazují uspokojivé výsledky. V u archivů se ale tento typ analýzy používá méně zejména proto, že esence v rozsáhlých archivech není pro prohlížení dostupná v plné kvalitě, ale pouze jako náhled nebo jen v podobě referenčních metadat. Vzhledem k tomu, že stále více digitálních úložišť je dostupných po síti, vznikají komerčně dostupná řešení (např. Blinx) určená k vyhledávání a indexování obecného audiovizuálního materiálu. Analýza obsahu zde není zaměřená na vytváření metadat navázaných na časový kód, ale na obecnou identifikaci typu obsahu v souboru a jeho přiřazení k určité kategorii nebo klíčovému slovu.

Specifickým případem jsou tzv. vizuální deskriptory, jak je definuje například standard MPEG-7 (Visual Descriptors). Jedná se o specifické atributy popisující strukturu, barevnost, pohyb a dokonce i standardizovanou charakteristiku tváří obrazu. Tyto atributy získané analýzou obrazu lze pak uložit i ve formátu XML s vazbou na popisovaný materiál. Jelikož příslušné deskriptory mohou být na rozdíl od archivované esenci přístupné uživateli okamžitě, dají se efektivněji použít pro vyhledávání. Tento typ analýzy lze použít i pro generování klíčových snímků. Otázkou zůstává exportní formát pro takto získané atributy, jelikož například v případě MPEG-7 nebo kompletního P/Meta souboru je mapování do MXF přímým způsobem neproveditelné. Lze využít ovšem kódování XML (MPEG-7) resp. HTML (P/META) pro připojení určitých atributů. Přes unikátní identifikátory lze ovšem zajistit pevnou vazbu na dodatečná metadata umístěná v databázích.

Systém DAM

Jak již bylo dříve zmíněno, u produkčních systémů je archiv úzce navázán na výrobní řetězec a systém DAM plní jak archivní, tak produkční funkci. Výrobní řetězec může být lineární, což odpovídá v dnešní době stále méně časté jednorázové produkci jednoho typu obsahu. Výrobní řetězec s recyklací materiálu (v angličtině vhodnější „repurposing“) používá výsledný materiál jako zdroj pro další operace, proto musí systém DAM tuto situaci zohlednit (Austerberry, 2006).



Příklad lineárního řetězce a řetězce s recyklací materiálu

Ačkoli řadu operací spojených s výrobou a zpracováním audiovizuálního materiálu zajišťují jiné systémy, systém DAM bývá z funkčních důvodů s řadou operací propojen. Typické operace, které systém DAM zajišťuje, nebo se jich účastní, jsou následující:

- Vstup audiovizuálního materiálu nebo obsahu (ingest)
- Vytváření náhledu, případně kódování do cílového (archivního) formátu
- Vstup a vytváření metadat
- Zařazení metadat do databáze
- Vstup a přiřazení autorských práv k databázovým položkám
- Ukládání obsahu na zvolené nosiče
- Vyhledávání v databázi
- Aktualizaci metadat nebo práv
- Export obsahu nebo jeho části z archivu (retrieval)
- Migraci obsahu na nové nosiče
- Migraci a zálohování databáze

Pro ukládání a přístup k esenci je třeba zajistit technologii, která dokáže nekódovat resp. dekódovat data do/z použitého formátu a tato data zobrazit ve vnímatelné formě. Systém DAM je z koncepčního pohledu na konkrétní kódovací a vizualizační technologii nezávislý, pracuje zejména s textovými popisy a náhledy a lze na něj nahlížet jako na informační systém. Zároveň do něj však řada informací (metadata, práva) vstupuje zvenčí v předem daném formátu a v definovaných návaznostech. Tyto návaznosti jsou diktovány konkrétním typem výroby a použitými technologiemi, které mají často deterministický vliv na podobu a obsah (zejména technických) metadat.

V praxi mohou být také funkce systému DAM rozděleny mezi bloky zajišťující specifické operace, nejvíce se tento princip uplatňuje při SOA (Service Oriented Architecture). Jednotlivé bloky tak mohou mít vlastní databáze a vlastní struktury metadat, typicky se jedná například o redakční systém nebo programový plánovací systém. Řada metadat tak existuje pouze jako součást specifických aplikací, mezi nimiž se přenáší pouze v případě, že jsou třeba pro nějakou produkční funkci. Pro jejich přenos se v tomto se používají proprietární formáty, XML, nebo nejnověji standard BXF. Z hlediska fungování celého uzavřeného systému tato situace nepředstavuje problém do okamžiku, kdy je potřeba vyměnit jeho rozsáhlou část nebo obsah exportovat pro nový typ použití (například pro web).

Doporučení pro systémy DAM a archivaci

Z hlediska dlouhodobé archivace je třeba zajistit, aby se primární metadata stala součástí archivovaného obsahu. Pro instituce, které se přímo výrobou audiovizuálního obsahu to znamená nejen zajištění přísunu těchto metadat, ale i schopnost pracovat se s formáty specifickými v audiovizuálním průmyslu, například s výše uvedeným formátem MXF. S tím se pojí i potenciální rozšíření nebo úprava datového modelu tak, aby umožňoval mapování například ze standardu MPEG-7 nebo lépe MPEG-21. Z hlediska dlouhodobého archivu již existují studie, které popisují mapování tohoto standardu na informační model OAIS (Bekaert, 2006). V praxi není nutné, aby pro přístup ke všem typům obsahu existoval jednotný DAM, protože požadavky na správu a přístup ke specifickým datovým strukturám se mohou značně lišit. Zde lze výhodou uplatnit přístup s využitím centrálního portálu, který dovoluje jednotný přístup k jednotlivým systémům pro správu na úrovni metadat a pro potřeby práce se samotným obsahem již volá klienta příslušného subsystému.

Použitá literatura:

AUSTERBERRY, D.: *Digital Asset Management, Second Edition*. 2 Focal Press, September 2006. ISBN 0-240-80868-1

BEKAERT, J., DE KOONING, E., VAN DE SOMPEL, H.: *Representing digital assets using MPEG-21 Digital Item Declaration* In: *Int. J. on Digital Libraries* 6(2): 159-173 Springer, 2006

DE SUTTER, R, STIJN NOTEBAERT, AND RIK VAN DE WALLE: *Evaluation of metadata standards in the context of digital audio-visual libraries* In: *LECTURE NOTES IN COMPUTER SCIENCE*, Springer 2006, ISSN 0302-9743

GRUBER, T.: *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. In.: *International Journal Human-Computer Studies* Vol. 43, Issues 5-6, November 1995, p.907-928.

KAHN, R. E.; WILENSKY, R. A.: *Framework for Distributed Digital Object Services*. In.: *International Journal on Digital Libraries* . 2006, vol. 6, no. 2, s. 115-123. Poprvé publikováno 1995-05-13

LEE, K., SLATTERY, O. ,LU, R.,TANG, X. ,McCRARY,V.: *The State of the Art and Practice in Digital Preservation*. In.: *Journal of Research of the National Institute of Standards and Technology*, 107, č.1, 2002. nvl.nist.gov/pub/nistpubs/jres/107/1/j71lee.pdf Shlédnuto 21.12.2008.

MEYROWITZ, J. : *Všude a nikde: vliv elektronických médií na sociální chování*, 1. vyd. Praha , Karolinum, 2006. ISBN: 80-246-0905-3

SAWYER, D. , REICH, L., GIARETTA, D., MAZAL, P. , HUC, C., NONON-LATAPIE, M., PECCIA, N.: *The Open Archival Information System (OAIS) Reference Model and its usage*. In: *SpaceOps2002*. 2002 Houston, TX. 9–12 Oct.

WELLS, N., MORGAN, O., WILKINSON, J., DEVLIN, B.: *The MXF Book: An Introduction to the Material eXchange Format*, Focal Press; 1st edition, 2006, ISBN: 978-0240806938