
Dlouhodobá ochrana elektronických publikací

Digital Preservation od Electronic Publications

Zuzana Kvašová

Národní knihovna České republiky

Tomáš Svoboda

Národní knihovna České republiky

Recenzenti:

Prof. Ing. Milan Konvit, PhD.

Prof. PhDr. Jaroslav Šušol, PhD.

Abstrakt:

Elektronické publikace se stávají předmětem zájmu ve stále širším měřítku. V knihovnách se řeší jejich sběr a zpřístupnění uživatelům. V rámci projektu Správa elektronických publikací v síti knihoven ČR má být zajištěna nejen potřebná infrastruktura pro sběr elektronických publikací a tiskových předloh, ale bude řešeno i dlouhodobé uchování těchto dokumentů. Tzv. Long Term Preservation je hlavním tématem tohoto článku, který se zabývá jak ochranou elektronických publikací v širším měřítku, tak konkrétními možnostmi uložení do LTP systému Národní knihovny ČR.

Klíčová slova: dlouhodobá ochrana digitálních dokumentů, elektronické publikace, dlouhodobá archivace, OAIS, metadata, formáty digitálních dat, metadatové standardy

Abstract:

Electronic publications are attracting a growing level of widespread interest. Libraries have to deal with their acquisition and user access to these. Goals of the project of „Electronic Publications Management in the Czech Republic Library Network“ are to design and implement a system of acquisition of electronic publications and pre-prints documents as well as digital preservation of these materials. Long term preservation is the main topic of the present article which focuses both on digital preservation in general and the particular solution of depositing these publications to a trusted repository of National Library of the Czech Republic.

Keywords: digital preservation, electronic publications, long term preservation, OAIS, metadata, digital formats, metadata standards

Práce s elektronickými publikacemi v Národní knihovně

Už několik let platí jednoznačný požadavek, aby Národní knihovna a další knihovny v ČR byly schopny přijímat do svých fondů publikace v digitální podobě, a to ve dvou podobách:

- dokumenty určené ke konzumaci elektronickou cestou (e-knihy) a

- předtiskové obrazy papírových dokumentů (preprinty)

Motivace vychází jednak z knihovnické komunity, která si uvědomuje, že touto cestou lze zachovat nezanedbatelnou část kulturního bohatství a také ze strany vydavatelů, kteří mají zájem mít své produkty bezpečně uschované a k dispozici v síti knihoven. Na základě těchto požadavků byl v rámci dotačního programu NAKI ministerstva kultury koncipován projekt „Správa elektronických publikací v síti knihoven ČR“ rozvržený na léta 2012 – 2015 (dále jen „projekt NAKI“). Jeho hlavním cílem je vytvořit poloprovoz, ve kterém budou vydavatelé odevzdávat a knihovny přijímat elektronické publikace.

Obě dvě tyto motivace přímo souvisí s problematikou dlouhodobého bezpečného uložení a udržování digitálních dokumentů, které je věnován tento článek.

Hlavní odlišnosti od práce s papírovými dokumenty jsou tyto:

- Jedna publikace se zpravidla vyskytuje ve více formátech s různými ISBN. Projekt musí vyřešit, jak zacházet s těmito vícenásobnými obrazy, které nejsou zcela totožné, ale mají mezi sebou tak silný vztah, že je třeba s nimi zacházet jako s jedinou významovou entitou.
- Digitální ochrana elektronických publikací má ještě větší důležitost, než u zdigitalizovaných papírových knih, protože neexistuje hmotná podoba, ke které by bylo možno se vrátit.

Pracovní postup při odevzdávání elektronické publikace lze zjednodušeně popsat takto: Producent, již dříve zaregistrovaný, do systému vloží datový soubor (EPUB2, PDF/A, případně další) a doplní k němu popisná metadata (doplnění nebude jen ruční, ale použije metadata z automatizovaných zdrojů). Systém provede kontrolu validity souboru, bezpečnostní kontrolu a kontrolu úplnosti metadat. Po úspěšných kontrolách se z dat i metadat vytvoří balíček SIP, který je vložen do systému dlouhodobé ochrany – LTP. Z tohoto balíčku se odvodí další deriváty, zejména pro zpřístupnění.

Odevzdávání elektronických publikací je v současné době dobrovolné. Odměnou za tuto dobrovolnost bude producentům dlouhodobé bezpečné uložení a garantované sdílení s třetími osobami podle rozhodnutí producenta, který určuje pravidla užití svého dokumentu.

První provozní pokusy s vybranými producenty začnou probíhat ve 2. pololetí 2013 a pilotní provoz začne o rok později, aby byl od roku 2015 nahrazen plným provozem příjmu a ochrany elektronických publikací, který budou využívat i další oprávněné knihovny v ČR.

Dlouhodobá ochrana digitálních dokumentů

Cílem dlouhodobé ochrany digitálních dokumentů neboli dlouhodobé archivace (angl. digital preservation) je zajištění trvalé udržitelnosti digitálních dat. Tím se rozumí především trvalý přístup k informacím, nezávislý na degradaci nosičů dat a technologických změnách, a udržení jejich srozumitelnosti a využitelnosti.¹ Digitální objekty by měly být ochráněny před zničením a ztrátou důvěryhodnosti a to v dlouhodobém horizontu.

¹ CUBR, Ladislav. Dlouhodobá ochrana digitálních dokumentů. 1. vyd. Praha : Národní knihovna České republiky, 2010. 154 s. ISBN 978-80- 7050-588-5 (brož.).

Dlouhodobou ochranu je možné rozdělit na tzv. ochranu bitů (bit-stream preservation) a logickou ochranu (content preservation), která spočívá v uchování informačního obsahu. Digitální objekty jsou uchovávány na nosičích, které podléhají degradaci. Samotná dlouhodobá ochrana tak musí začínat u bitové ochrany, která se nejčastěji realizuje redundantním uložením a řízeným kopírováním na nové nosiče. Součástí logické ochrany jsou procesy, které mají zajistit ochranu obsahu tak, aby význam informací zůstal zachován. Je nutné se vypořádat se zastaráváním softwarových aplikací, které jsou využívány pro zpřístupnění obsahu. Toto je řešeno formátovou migrací, či emulací. Opatření digitálních objektů vhodnými metadaty vede k zajištění autenticity a důvěryhodnosti obsahu. Metadaty mají popsat obsah, strukturu a kontext informačního objektu.² Aktivita v oblasti dlouhodobé ochrany dat popisuje referenční model OAIS (Open Archival Information System), schválený jako mezinárodní norma ISO 14721:2003, resp. v aktualizaci ISO 14721:2012. Podle tohoto modelu by digitální informace, udržované v rámci dlouhodobé archivace, měly být spravovány tak, aby bylo možné se vypořádat s měnícími se technologiemi a změnami očekávání uživatelské komunity. Jejich správu zajišťuje digitální archiv, ve kterém jsou uloženy. Norma OAIS definuje tento archiv jako organizaci lidí a systémů se závazkem ochraňovat a zpřístupňovat digitální data pro určitou skupinu uživatelů.³

Model OAIS je možné využít rovněž pro dlouhodobou ochranu elektronických publikací. V rámci projektu NAKI vzniká návrh procesu ukládání elektronických publikací do LTP systému. Počítá se se sběrem a ukládáním jak elektronických předloh tištěných verzí publikací, tak samotných e-publikací.

Elektronické publikace a jejich dlouhodobá archivace

Elektronické publikace mohou obsahovat různý typ obsahu. Pro dlouhodobou ochranu je nutné zajistit především uchování textu tak, aby byla zachována srozumitelnost.

V závislosti na typu dokumentu je navíc třeba rozlišit, nakolik je kromě textového obsahu vhodné uchovávat i vizuální podobu, ilustrace a netextové objekty. Například v odborné literatuře může být důležité zachovat tabulky, rovnice apod., v umělecké publikaci i celkový design dokumentu. Naproti tomu u běžné beletrie bude pravděpodobně pro zachování informačního obsahu stačit uchovávat text.

Pro potřeby dlouhodobé ochrany bude potřeba uchovat i další vlastnosti dokumentu:

- „kvantitativní informace o textu dokumentu, jako je počet znaků, případně další informace, počet odstavců,
- informace související s kódováním textu, jazykem,
- strukturální informace, které umožní text číst tak, jak bylo zamýšleno (informace o navigaci dokumentem),

² HUTAŘ, Jan. Digitalizace, popis pomocí metadat a jejich formáty. Praha, 2012. 244 s. Disertační práce. Univerzita Karlova v Praze, Ústav informačních studií a knihovnictví. Dostupný také z: <https://is.cuni.cz/webapps/UKSESSION07FE21533837C5377A6362EE8AE949EE/zzp/detail/25756/?lang=cs&file=140015545>.

³ ROSENTHAL, Colin, Asger BLEKINGE-RASMUSSEN a Jan HUTAŘ. Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER). 1. vyd. Praha: Národní knihovna České republiky, 2009, 51 s. ISBN 978-807-0505-694.

- informace o původu dokumentu (kdo ho publikoval, identifikátory, bibliografická metadata) a technické informace o tvůrci dokumentu a použitém SW,
- ne-textové vizuální objekty (jako diagramy, rovnice, ilustrace) a informace o jejich umístění v textu,
- technické informace o digitálním formátu,
- popisná metadata k obsahu,
- kontextové informace, vazby na jiné dokumenty a objekty⁴

Digitální formáty pro dlouhodobou ochranu

Určení vhodného formátu pro dlouhodobou ochranu má dopad na budoucí možnosti uchování elektronických publikací a obsahových informací v použitelné podobě.

Podle studie, která vznikla v Kongresové knihovně, je možné určit sedm faktorů udržitelnosti a použít je jako kritéria vhodnosti formátu.⁵ Těmito faktory jsou:

- Otevřenost – dostupnost specifikace a validačních nástrojů. Důležité pro porozumění způsobu reprezentace informací v souboru.
- Rozšířenost – nakolik je formát využíván tvůrci obsahu, distributory a uživateli informačních zdrojů.
- Transparentnost – otevřenost formátu analýzám různými nástroji třetích stran.
- Sebeopisnost – možnost popsání metadata přímo v digitálním objektu, především technickými metadaty za účelem udržení informace potřebné k použití formátu.
- Nezávislost – stupeň závislosti na konkrétním hardware, operačním systému či software.
- Dopad patentů – možné zvýšení nákladů na zachování obsahu.
- Mechanismy technické ochrany – obsah určený pro dlouhodobou ochranu nesmí mít omezenou přístupnost, např. DRM ochranu.

V rámci projektu NAKI byly určeny nejčastěji využívané formáty elektronických publikací na českém trhu – PDF, EPUB, MOBI, WOOLY, IBOOK. Dalšími formáty elektronických knih jsou AZW, AZW1, AZW4, KF8, PDB, TXT, DJVU, BBeB Book, Microsoft LIT a FB2. Vybraným formátům byly věnovány samostatné studie.

Cílem bylo vybrat vhodné formáty tak, aby splňovaly kritéria nejčastěji využívaného formátu a formátu vhodného pro dlouhodobou ochranu.

Proprietární formáty IBOOK a WOOLY byly deklarovány jako nevhodné pro potřeby dlouhodobé ochrany, navíc jsou opatřeny proprietárními DRM technologiemi (více text níže).

Formát MOBI je formátem částečně proprietárním, není kompletně zveřejněna jeho dokumentace. V rámci analýzy byl, pro svou podobnost, porovnáván s formátem EPUB. Oba formáty dokáží upravit tok textu v závislosti na velikosti a rozlišení displeje čtecího zařízení, umožňují nastavení

⁴ MELICHAR, Marek. Technické analýzy formátů elektronických dokumentů. Praha: Národní knihovna ČR, 30. 12. 2012. Nepublikovaný dokument.

⁵ LIBRARY OF CONGRESS. Sustainability of Digital Formats: Planning for Library of Congress Collections [online]. Aktualizace 2013 [cit. 2013-06-18]. Sustainability Factors. Dostupný z WWW: <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.

velikosti písma podle preferencí uživatele a v omezené míře umožňují ukládání multimediálního obsahu do elektronických publikací. Žádný z nich nepodporuje stránkování a poskytují omezené možnosti kontroly nad rozvržením obsahu e-publikace.

Formát EPUB má oproti formátu MOBI otevřenou specifikaci. Umožňuje uchovat i multimediální obsah, verze 3.0 i graficky náročnější publikace. Verze EPUB 2.0 graficky náročné publikace uchovat neumí, ovšem jeho velká rozšířenost, a tím i škála existujících editačních a validačních nástrojů, je prozatím pro potřeby dlouhodobé ochrany vhodnější.

Analýze vlastností pro dlouhodobou archivaci byl podroben i formát DJVU, který byl v minulosti hlavním konkurentem formátu PDF. Oba formáty mohou obsahovat vícestránkový dokument s obrázky a textem v rámci jednoho souboru.⁶ Formát DJVU je však částečně proprietární a není příliš podporován jeho rozvoj.

PDF formát je souborový formát, který vznikl v roce 1991 a vyvíjela ho společnost Adobe. V roce 2008 se stal mezinárodním standardem. Společnost Adobe Systems Incorporated formát stále vyvíjí, avšak záštitu převzala technická komise ISO. Jedná se o rozšířený formát, který podporuje složené dokumenty – text i grafiku v rámci jednoho souboru. Pracuje s kompresí tak, aby byla zachována čitelnost textu i při malých souborech, podporuje plnotextové vyhledávání a umožňuje jazykové mutace v rámci jednoho souboru. K zobrazování i editaci je třeba specializovaného software či pluginů. Vznikla jeho archivní podoba PDF/A1 a PDF/A2. PDF/A je kontejner určený především k archivaci a pro účely dlouhodobé ochrany. Je upravován mezinárodními standardy ISO 19005-1:2005 a ISO 19005-2:2011. Důraz je kladen na otevřenou specifikaci a sebedopisnost – povinné vkládání metadat umožňujících vlastní identifikaci a vkládání historie v celém procesu. Především z těchto důvodů je pro potřeby dlouhodobé ochrany vhodný.

Na základě analýz bylo rozhodnuto, že pro uložení v systému NK pro dlouhodobou ochranu budou využívány formáty EPUB verze 2 a PDF, respektive PDF/A. Ostatní formáty budou sbírány a ukládány, ale bude nad nimi zajišťována pouze bitová ochrana.

Tato volba vycházela z výsledků, ze kterých vyplynulo, že většina elektronických publikací je vydávána ve formátu PDF, který dominuje především u elektronických časopisů a tiskových předloh. Elektronické knihy jsou navíc často publikovány ve formátech EPUB a MOBI. V rámci projektu bude vytvořen software pro převod preferovaných formátů do formátů vhodných pro uložení v LTP systému (systém pro dlouhodobou archivaci) a následnou dlouhodobou ochranu. Původní formát dat zůstane v LTP systému zachován a bude nad ním realizována bitová ochrana.

Elektronické publikace často obsahují DRM (Digital Rights Management) zabezpečení, jehož cílem je zamezit zneužití obsahu digitálních dokumentů.⁷ Nejčastěji se jedná o Adobe DRM, využívanou pro formáty EPUB a PDF, která funguje na bázi autorizace uživatele, nebo o tzv. sociální DRM, kdy je publikace opatřena osobními údaji uživatele, který ji zakoupí. Analyzovaný formát IBOOK využívá technologii FairPlay DRM, formát WOOLY je opatřen Wooly DRM. Sociální DRM by z hlediska digitální archivace nebyla problémem, ovšem využití jakéhokoliv šifrování by pro

⁶ VYCHODIL, Bedřich. *Formátová studie*. Praha, Národní knihovna, 5. 10. 2012. Nепublikovaný dokument.

⁷ PUCKETT, Jason. Digital Rights Management as Information Access Barrier. *Georgia State University: Digital Archive @GSU* [online]. 2010 [cit. 19-06-2013]. Dostupný z WWW:

http://digitalarchive.gsu.edu/cgi/viewcontent.cgi?article=1049&context=univ_lib_facpub

dlouhodobou ochranu překážkou bylo. Bylo rozhodnuto, že v rámci sběru elektronických publikací nebudou přijímány dokumenty s DRM zabezpečením.

Validace digitálních formátů a extrakce technických metadat

V rámci projektu byla provedena analýza dostupných nástrojů pro validaci dat.⁸ Nad daty ve formátu EPUB verze 2 bylo otestováno 6 existujících nástrojů. K identifikaci formátu objektů ve formátu EPUB byl použit Droid verze 6, signatures 64, který se osvědčil. Soubory bez DRM je schopen identifikovat podrobně, včetně obsahu. Pro validaci obsahu ve formátu EPUB byly použity nástroje Epubcheck a FlightCrew. Pro validaci rozbaleného obsahu byl využit nástroj Jhove verze 1.8. Pro identifikaci formátů byly otestovány i nástroje Apache Tika a File, jako alternativa k nástroji Droid. Z analýzy vyplynulo, že nemá smysl je zapojovat společně s nástrojem Epubcheck, který navíc dodává i výsledek validace. Vzhledem k tomu, že Flightcrew není dále aktivně podporován uživatelskou komunitou, bude nutné využívat pro validaci formátů epub Epubcheck. Pro validaci formátu PDF/A se ukázalo jako nejlepší využití komerčního validátoru Adobe Reader, který kontroluje shodu s ISO normou v plném rozsahu. Do budoucna bude možné sledovat rozvíjení nástroje PDFBOX, který se prozatím neosvědčil z důvodu nestability software a problémy s přerušením zpracování po detekci odchylky od specifikace.⁹ K extrakci technických metadat může být využit nástroj Jhove.

V rámci projektu NAKI bude o využití konkrétních nástrojů pro validaci a extrakci technických metadat rozhodnuto v budoucnu.

Metadata v archivním balíčku a archivní balíček

Pro uložení elektronických publikací v LTP systému NK bude vytvořen návrh metadatové specifikace, která bude určovat podobu AIP balíčku.

Podle standardu OAIS je pro dlouhodobou ochranu nezbytný popis metadaty. Ta jsou v současnosti dělena nejčastěji na metadata popisná, administrativní, kam spadají metadata práv, ochranná metadata a technická metadat a na metadata strukturální.¹⁰

Popisná metadata popisují digitální objekt, jako intelektuální entitu z bibliografického pohledu. Tato metadata se využívají především pro vyhledání a identifikaci digitálního obsahu. Obsahují údaje o autorovi, názvu, vydavateli, identifikátory a další. Kromě formálního popisu mohou

⁸ MELICHAR, Marek. *Technické analýzy formátů elektronických dokumentů*. Praha: Národní knihovna ČR, 30. 12. 2012. Nepublikovaný dokument.

⁹ Van der KNIJFF, Johann. Identification on PDF preservation risks with Apache Preflight: a first impression. Open Planets Foundation. Open Planets Foundation: A community hub for digital preservation [online]. 19. 12. 2012 [cit. 20-06-2013]. Dostupný z WWW: <http://www.openplanetsfoundation.org/blogs/2012-12-19-identification-pdf-preservation-risks-apache-preflight-first-impression>

¹⁰ HUTAŘ, Jan. *Digitalizace, popis pomocí metadat a jejich formáty*. Praha, 2012. 244 s. Disertační práce. Univerzita Karlova v Praze, Ústav informačních studií a knihovnictví. Dostupný také z: <https://is.cuni.cz/webapps/UKSESSION07FE21533837C5377A6362EE8AE949EE/zp/detail/25756/?lang=cs&file=140015545>

zahrnovat i věcný popis. Výchozím bodem pro tento druh metadat v rámci projektu NAKI bude výstup z analýzy bibliografického popisu, který vznikne v rámci projektu. Tato metadata budou pravděpodobně přebírána po zkatalogizování publikace ze systému Aleph. Dodavatelé budou elektronické publikace opatřovat svými metadaty, která umožní identifikaci dané publikace. Schéma popisných metadat bude zvoleno, je pravděpodobné, že se bude jednat o schéma MODS či Dublin Core.

Administrativní metadata jsou určena ke správě dat, slouží především administrátorovi systému. Obsahují informace o původu intelektuální entity, informace o akvizici, o právech souvisejících s intelektuální entitou, údaje o lokaci. Rovněž obsahují podmnožinu technických metadat s informacemi o technických vlastnostech jednotlivých digitálních souborů a jejich formátů, o technickém prostředí využitém při vytvoření a nutném pro zobrazení a kontrolní součty. Je pravděpodobné, že pro technická metadata bude použit standard PREMIS. Ochranná metadata nesou informace, které jsou potřeba pro dlouhodobou ochranu. Patří mezi ně informace o akcích provedených nad uloženým objektem, např. migrace či aktualizace.

Pro popsání struktury a vztahy mezi digitálními objekty jsou využívána strukturální metadata. Ta popisují buď fyzickou strukturu objektu (případně objektů), vyjadřující, které části digitálních dat mají vzájemnou vazbu, nebo logickou strukturu. Ta vyjadřuje logickou stavbu dokumentu, např. kapitoly, články.¹¹ Také jejich zachování může být klíčové pro pochopení dokumentu čtenářem v daleké, nedefinované budoucnosti. V rámci projektu bude pravděpodobně využito formátu METS, který je nyní užíván i v metadatovém standardu pro zdigitalizované dokumenty.

Závěr

Dlouhodobá ochrana elektronických dokumentů je v Národní knihovně brána velmi vážně a tento přístup se projevuje i v projektu „NAKI - Správa elektronických publikací v síti knihoven ČR“. Příjem elektronických publikací je proto navrhován se zřetelem na bezpečné uložení těchto dokumentů a jejich uchování pro budoucí generace. Můžeme proto oprávněně očekávat, že i tato část kulturního bohatství, která dnes již vzniká v elektronické podobě, bude zachována v Národní knihovně nejméně tak kvalitně, jako dokumenty z minulých staletí.

Literatura:

1. PDF Technology Center. ADOBE SYSTEMS INCORPORATED. *Adobe* [online]. 2013 [cit. 2013-06-19]. Dostupný z WWW: <http://www.adobe.com/devnet/pdf.html>
2. CUBR, Ladislav. *Dlouhodobá ochrana digitálních dokumentů*. 1. vyd. Praha : Národní knihovna České republiky, 2010. 154 s. ISBN 978-80-7050-588-5 (brož.).
3. HUTAŘ, Jan. *Digitalizace, popis pomocí metadat a jejich formáty*. Praha, 2012. 244 s. Disertační práce. Univerzita Karlova v Praze, Ústav informačních studií a knihovnictví. Dostupný z WWW:

¹¹ HUTAŘ, Jan. *Digitalizace, popis pomocí metadat a jejich formáty*. Praha, 2012. 244 s. Disertační práce. Univerzita Karlova v Praze, Ústav informačních studií a knihovnictví. Dostupný také z: <https://is.cuni.cz/webapps/UKSESSION07FE21533837C5377A6362EE8AE949EE/zp/detail/25756/?lang=cs&file=140015545>

- <https://is.cuni.cz/webapps/UKSESSION07FE21533837C5377A6362EE8AE949EE/zp/detail/25756/?lang=cs&file=140015545>.
4. KNIJFF, Johann van der. Identification on PDF preservation risks with Apache Preflight: a first impression. Open Planets Foundation. *Open Planets Foundation: A community hub for digital preservation* [online]. 19. 12. 2012 [cit. 20-06-2013]. Dostupný z WWW: <http://www.openplanetsfoundation.org/blogs/2012-12-19-identification-pdf-preservation-risks-apache-preflight-first-impression>.
 5. KNIJFF, Johann van der. EPUB for archival preservation: an update. Open Planets Foundation. *Open Planets Foundation: A community hub for digital preservation* [online]. 23. 05. 2013 [cit. 20-06-2013]. Dostupný z WWW: <http://www.openplanetsfoundation.org/blogs/2013-05-23-epub-archival-preservation-update>.
 6. LIBRARY OF CONGRESS. *Sustainability of Digital Formats: Planning for Library of Congress Collections* [online]. Aktualizace 2013 [cit. 2013-06-18]. Sustainability Factors. Dostupný z WWW: <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.
 7. MELICHAR, Marek. *Technické analýzy formátů elektronických dokumentů*. Praha: Národní knihovna ČR, 30. 12. 2012. Nepublikovaný dokument. Vznik v rámci projektu Správa elektronických publikací v síti knihoven ČR.
 8. PUCKETT, Jason. Digital Rights Management as Information Access Barrier. *Georgia State University: Digital Archive @GSU* [online]. 2010 [cit. 19-06-2013]. Dostupný z WWW: http://digitalarchive.gsu.edu/cgi/viewcontent.cgi?article=1049&context=univ_lib_facpub
 9. ROSENTHAL, Colin, Asger BLEKINGE-RASMUSSEN a Jan HUTAŘ. *Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER)*. 1. vyd. Praha: Národní knihovna České republiky, 2009, 51 s. ISBN 978-807-0505-694.
 10. VYCHODIL, Bedřich. *Formátová studie*. Praha, Národní knihovna, 5. 10. 2012. 23 s. Nepublikovaný dokument. Vznik v rámci projektu Správa elektronických publikací v síti knihoven ČR.