

*CZech IDentification and LOcazation tool / Český systém pro IDentifikaci a LOkalizaci
digitálních dokumentů českého kulturního dědictví*

Metodika pro přidělování a správu životního cyklu unikátních perzistentních identifikátorů digitálních dokumentů podle standardu URN:NBN

Zdeněk Vašek, Ladislav Cubr a kol.

*Realizováno v rámci institucionálního výzkumu Národní knihovny České republiky v části „2b
Vývoj podpůrných/návazných aplikací pro dlouhodobou ochranu digitálních dat“
financovaného Ministerstvem kultury ČR v rámci Dlouhodobého koncepčního rozvoje
výzkumné organizace*

Perzistentní identifikátor dokumentu: urn:nbn:cz:nk-0027gj

Obsah

Seznam obrázků a tabulek.....	4
1. ÚČEL METODIKY	5
2. URČENÍ.....	6
3. SOUVISEJÍCÍ NORMY A PŘEDPISY	7
4. TERMINOLOGIE.....	10
5. IDENTIFIKÁTORY.....	15
5.1 Přehled stavu - identifikátory v praxi knihoven.....	15
5.1.1 Identifikátory tištěných dokumentů	15
5.1.2 Identifikátory a digitální dokumenty.....	17
5.2 Identifikátor URN:NBN	21
5.3 Východiska systému ČIDLO.....	22
6. SYSTÉM ČIDLO	25
6.1 Resolver.....	25
7. PODMÍNKY ÚČASTI V SYSTÉMU ČIDLO	27
8. MODEL ROLÍ A ODPOVĚDNOSTÍ V SYSTÉMU ČIDLO.....	29
8.1 Návod pro zápis URN:NBN identifikátorů do elektronického katalogu	31
9. PŘEDMĚT IDENTIFIKACE.....	32
9.1 Model dokumentu v systému ČIDLO	33
9.1.1 Intelektuální entita.....	33
9.1.2 Digitální reprezentace/ dokument	35
9.1.3 Digitální instance.....	35
9.1.4 Vzájemné vztahy	36
9.1.5 Datový model	36
9.2 Vznik a struktura digitalizovaného dokumentu.....	38
10. POSTUP PŘIDĚLOVÁNÍ IDENTIFIKÁTORŮ, PRAVIDLA SYSTÉMU ČIDLO ...	40
10.1 Obecné povinnosti.....	40

10.2 Povinnosti před přidělením URN:NBN identifikátoru (registrace registrátora)	41
10.3 Uživatelské účty	42
10.3.1 Možnosti operací při správě jmenného prostoru	42
10.4 Postup ve fázi přidělení URN:NBN identifikátoru	43
10.4.1 Příklady operací při registraci digitálního dokumentu	44
10.5 Povinnosti z hlediska zpřístupňování dokumentu	49
10.5.1 Příklad operace vložení digitální instance	49
10.6 Chybové stavy	51
11. TYPY DOKUMENTŮ	53
11.1 Pokyny pro registrace jednotlivých typů dokumentů	54
11.1.1 MONOGRAPH	54
11.1.2 MONOGRAPH_VOLUME	62
11.1.3 PERIODICAL_ISSUE	64
11.1.4 PERIODICAL_VOLUME	67
11.1.5 ANALYTICAL	68
11.1.6 DIGITAL DOCUMENT	69
11.2 Technická metadata zasílaná systému ČIDLO	72
12. RESOLVOVACÍ FUNKCE SYSTÉMU ČIDLO	75
13. HROMADNÉ SKLÍZENÍ POMOCÍ OAI-PMH	77
13.1 OAI Adapter	78
14. PRAVIDLA PRO ŽIVOTNÍ CYKLUS DOKUMENTU A IDENTIFIKÁTORU	79
14.1 Návod na postupy při slučování a rozpojování dokumentů	82
O AUTORECH	84
DOPORUČENÁ LITERATURA	84

Seznam obrázků a tabulek

Obrázek 1: Příklad identifikátoru pro NLK (kapitola 5.2)

Obrázek 2: Příklad identifikátoru pro NK (5.2)

Obrázek 3: Popis vztahů mezi intelektuální entitou a digitální reprezentací (9.1.4)

Obrázek 4: Konceptní model objektů v systému ČIDLO (9.1.5)

Obrázek 5: Vztahy mezi tabulkami v databázi (9.1.5)

Tabulka 1: Role v systému ČIDLO (kapitola 8)

Tabulka 2: Přehled chybových kódů v „Response“ (10.6)

Tabulka 3: Doporučení na získání hodnoty „číslo ČNB“ z metadat (11.1.1)

Tabulka 4: Doporučení na získání hodnoty „ISBN“ z metadat (11.1.1)

Tabulka 5: Doporučení na získání hodnoty „jiný identifikátor“ z metadat (11.1.1)

Tabulka 6: Doporučení na získání hodnoty „druh dokumentu“ z metadat (11.1.1)

Tabulka 7: Doporučení na získání hodnoty „hlavní název“ z metadat (11.1.1)

Tabulka 8: Doporučení na získání hodnoty „autor“ z metadat (11.1.1)

Tabulka 9: Doporučení na získání hodnoty „korporace“ z metadat (11.1.1)

Tabulka 10: Doporučení na získání hodnoty „akce“ z metadat (11.1.1)

Tabulka 11: Doporučení na získání hodnoty „další původce“ z metadat (11.1.1)

Tabulka 12: Doporučení na získání hodnoty „místo vydání“ z metadat (11.1.1)

Tabulka 13: Doporučení na získání hodnoty „název nakladatele“ z metadat (11.1.1)

Tabulka 14: Doporučení na získání hodnoty „rok vydání“ z metadat (11.1.1)

Tabulka 15: Vzorové mapování metadat MODS na pole databáze ČIDLO (monografie) (11.1.1)

Tabulka 16: Vzorové mapování metadat MODS na pole databáze ČIDLO (svazek monografie) (11.1.2)

Tabulka 17: Doporučení na získání hodnoty „druh dokumentu“ z metadat (číslo periodika) (11.1.3)

Tabulka 18: Doporučení na získání hodnoty „ISSN“ z metadat (číslo periodika) (11.1.3)

Tabulka 19: Vzorové mapování metadat MODS na pole databáze ČIDLO (číslo periodika) (11.1.3)

Tabulka 20: Vzorové mapování metadat MODS na pole databáze ČIDLO (článek) (11.1.5)

Tabulka 21: Vzorové mapování metadat na pole databáze ČIDLO (digitální dokument) (11.1.6)

Tabulka 22: Doporučené mapování z metadat ve standardu MIX (11.2)

1. ÚČEL METODIKY

Metodika předkládá postup pro získání, správné použití a následné nakládání s unikátními perzistentními identifikátory digitálních dokumentů kulturní povahy ze sféry knihoven v České republice. Metodika se zabývá životním cyklem identifikátorů založených na standardu URN:NBN (Uniform Resource Name: National Bibliography Number – standard RFC 3188¹). Definuje potřebné kroky k získání identifikátoru URN:NBN v českém prostředí (služba ČIDLO), pravidla pro jeho přidělování a zásady pro následné dlouhodobé uložení a zpřístupnění digitálního dokumentu, kterému byl identifikátor URN:NBN přidělen.

Cílem metodiky je poskytnout návod producentům a správcům digitalizovaných dokumentů, jak správně získávat perzistentní identifikátory v systému ČIDLO a jak nakládat s identifikovanými dokumenty tak, aby byla dodržena pravidla standardu URN:NBN na národní i mezinárodní úrovni a zároveň bylo vyhověno požadavkům na autenticitu, dlouhodobé uložení, správu a také zpřístupnění digitálního dokumentu. Správné využití návodu této metodiky ve všech zmíněných oblastech (od přidělení po zpřístupnění) představuje též nutnou podmínku pro používání identifikátoru URN:NBN jako jednoho ze způsobů, jak citovat online dokumenty, což výrazně přispěje ke zvýšení důvěryhodnosti citování ve vědecké praxi, především v oblasti humanitních věd. Využití standardu URN:NBN podle předložené metodiky je také jedním ze způsobů zajištění autenticity dokumentů a zároveň slouží pro propojení při dlouhodobé správě digitálních dokumentů

Pravidla a postupy metodiky byly ověřeny v letech 2012–2014 v rámci provozu Národní digitální knihovny a ve spolupráci s dalšími producenty digitálních dokumentů, především z programu VISK7 a tzv. krajských digitalizací.

¹<https://www.ietf.org/rfc/rfc3188.txt> Návrh podala Finská národní knihovna prostřednictvím Juhy Hakaly jako doporučení Internetové společnosti (The Internet Society) v roce 2001. Citováno [2014-09-22].

2. URČENÍ

Metodika je určena všem uživatelům služby ČIDLO, kteří jsou v roli registrátora nebo archivátora, případně jsou touto rolí pověřeni, dále také správcům digitálních repozitářů a digitálních knihoven paměťových institucí. Potřeba jednoznačné a trvalé identifikace se ukazuje jako stále důležitější otázka. Globálnost, perzistence a standardizace (pravidla) identifikátoru zaručují jeho dlouhodobou užitečnost při správě a vyhledávání dokumentů. V běžné praxi by měli metodiku využívat externí producenti dat, pracovníci digitalizačních jednotek provozovaných paměťovými institucemi (především knihovnami) a další pracovníci spravující sbírky digitálních dokumentů za účelem archivace i zpřístupnění. Cílovou skupinou metodiky jsou knihovny všech úrovní nebo obdobné instituce (např. muzea bez registrované knihovny, ale s knihovně sbírkou) a externí producenti dat zajišťující digitalizaci.

Základní vymezení uživatelské komunity:

- krajské a ústřední knihovny,
- knihovny vědeckých ústavů,
- vysokoškolské knihovny,
- producenti digitalizovaných dat kulturní povahy ve sféře knihoven.

V širším vymezení má služba ČIDLO dva okruhy uživatelů. První tvoří ti, jimž je určena tato metodika, a druhý čtenáři, kteří službu ČIDLO mohou využívat jako cestu k zjištění aktuálního umístění identifikovaných dokumentů v internetové síti a jako možnost pro citování těchto dokumentů.

3. SOUVISEJÍCÍ NORMY A PŘEDPISY

Samotná metodika není všeobjímajícím doporučením jak nakládat s digitálními dokumenty. Definuje pouze zacházení s identifikátorem URN:NBN a to, jaká omezení musí jeho uživatelé respektovat. Mezi ně patří i korektní využívání standardů a norem vztahujících se k samotnému identifikátoru, k archivaci dokumentů a k vytváření metadat a některých dalších předpisů. Tyto standardy, normy a předpisy je nutné používat, pokud mají být využity všechny možnosti a doporučení předkládané metodikou.

3.1 Standard URN:NBN

Tento standard je základním předpisem pro syntax a užívání identifikátoru URN:NBN². Jeho pravidla jsou implementována do systému ČIDLO, dodržováním jeho pravidel je garantováno i plnění standardu pro URN:NBN.

3.2 Normy pro digitální repozitář

Normy ISO 14721:2012 (Open archival information system – Reference model; česká verze jako ČSN ISO 14721: Otevřený archivační informační systém – Referenční model) a ISO 16363:2012 (Audit and certification of trustworthy digital repositories; česká verze jako ČSN ISO 16363: Audit a certifikace důvěryhodných digitálních úložišť) jsou dvě základní normy v oblasti dlouhodobého uchovávání digitálních dokumentů. Norma ISO 14721:2012 specifikuje požadavky na funkce archivačního systému (archiv OAIS) pro dlouhodobé uchovávání digitálních dokumentů. Jde o druhé vydání normy (první bylo v roce 2003). Norma ISO 16363:2012 specifikuje takové požadavky na audit a certifikaci archivu OAIS, aby mohl být označen za důvěryhodný.

Obě normy vyšly v roce 2014 též v češtině. Jako oficiální české verze mezinárodních norem je vydal Úřad pro technickou normalizaci, metrologii a státní zkušebnictví (<http://www.unmz.cz/urad/unmz>).

Archivní verze digitálních dokumentů s URN:NBN identifikátory mohou být uloženy i mimo systémy splňující normu OAIS, ovšem nelze tento postup dlouhodobě doporučit.

3.3 Metadata Object Description Schema (MODS)

Metadatový standard MODS dostupný z <http://www.loc.gov/standards/mods/> [citováno 2014-09-23] je výchozím standardem pro tvorbu popisných metadat, která jsou

² Dostupný z <https://www.ietf.org/rfc/rfc3188.txt> [citováno 2014-09-23].

následně předávána podle pravidel metodiky. Metodika předpokládá, že pro tvorbu metadat, která budou předávána systému ČIDLO za účelem získání identifikátoru URN:NBN, budou využity verze standardu 3.2; 3.3; 3.4 a 3.5. Využití standardu je povinné, jiné než doporučené verze standardu je předem nutno konzultovat s kurátorem systému ČIDLO.

3.4 MIX (NISO Metadata for Images in XML)

Technická metadata pro obrazová data definuje norma ANSI/NISO Z39.87-2006. Jejich zápis do XML stanovuje standard MIX. MIX slouží k podrobnějšímu popisu obrazových dat, používá se k popisu formátu digitálních obrazů, jejich rozlišení, reálné velikosti, barevného modelu atd. Zaznamená také údaje o hardware, který digitální obraz vytvořil. Metodika předpokládá využití verze 2.0 standardu MIX. Využití standardu je doporučené. Standard je dostupný z <http://www.loc.gov/standards/mix/> [citováno 2014-09-23].

3.5 Standardy digitalizace Národní digitální knihovny

Metodika předpokládá, že digitální dokumenty identifikované pomocí URN:NBN identifikátoru budou vytvářeny nebo přetvářeny pomocí standardů digitalizace Národní digitální knihovny, a to v jakékoliv verzi standardů, přičemž se doporučuje využít vždy aktuálně platnou verzi dostupnou z webu NDK: <http://www.ndk.cz/standardy-digitalizace>. Použití Standardů digitalizace je touto metodikou doporučené, není povinné, pokud to nevyžadují jiné předpisy. Standardy digitalizace vydala Národní knihovna ČR.

3.6 REST (Representational State Transfer)

Jde o architekturu rozhraní, navrženou pro distribuované prostředí, které využívá i software CZIDLO. Byla představena v roce 2000 v disertační práci Roye Fieldinga, jednoho ze spoluautorů protokolu HTTP. Rozhraní REST je použitelné pro jednotný a snadný přístup ke zdrojům (resources). Základní metodou pro přístup ke zdrojům je získání zdroje – metoda GET. REST je architektura, která umožňuje přistupovat k datům na určitém místě pomocí standardních metod http.

Definice je dostupná z

http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm [citováno 2014-09-23].

Pro automatické přidělování identifikátorů URN:NBN podle této metodiky je používání protokolu REST nezbytné. Částečně jej lze nahradit manuálním přidělováním pomocí webového rozhraní.

3.7 The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

Tento protokol slouží k hromadnému sklizení metadatových záznamů různých repozitářů, jako jsou katalogy, digitální knihovny apod. Jde o aplikační protokol nad HTTP. Podrobnosti k protokolu dostupné z <http://www.openarchives.org/OAI/openarchivesprotocol.html> [citováno 2014-09-23]. Tato metodika předpokládá, že digitální knihovny, ve kterých budou zveřejňovány dokumenty identifikované pomocí identifikátoru URN:NBN, bude možno sklízet pomocí OAI-PMH protokolu nebo budou mít záměr tuto možnost zavést. Pokud budou dokumenty zveřejňovány ve více knihovnách, pak je nutné, aby k tomu byla uzpůsobena alespoň jedna z těchto knihoven.

3.8 Dublin Core (DC)

Metadatový standard Dublin Core dostupný z <http://dublincore.org/> [citováno 2014-09-23] je standardem pro tvorbu popisných metadat. Tato metodika ho chápe jako doplňující, nicméně je vyžadován dalšími předpisy.

4. TERMINOLOGIE

Zkratky a pojmy užívané v metodice

Tyto zkratky a pojmy vycházejí z praxe knihoven v ČR (jedná se především o vžité zkratky), z ustáleného názvosloví užívaného v rámci systému ČIDLO a dále z odborné terminologie, kterou je nutné v metodice využít pro správný popis jednotlivých kroků. Pojmy z normy OAIS jsou uváděny podle české verze normy.

ADR

Adresář knihoven dostupný z <http://www.nkp.cz/katalogy-a-db>

AIP

Archivní informační balíček (archival information package)

informační balíček, který je složen z informačního obsahu a přidružených informací o uchovávání a je uchováván v archivu OAIS

Aktuální archivní dokument

Archivní dokument, který je obsažen v nejnovější verzi balíčku AIP.

Archivátor

Instituce, která provozuje digitální repozitář.

Archivní dokument

Archivní kopie digitálního dokumentu uložená v digitálním repozitáři v balíčku AIP..

Bibliografická metadata

Popisná bibliografická metadata ve standardu MODS.

ČIDLO

(Český systém pro IDentifikaci a LOkalizaci digitálních dokumentů českého kulturního dědictví)

ČIDLO je systém pro trvalou identifikaci digitálních dokumentů českého kulturního dědictví.

Tento systém využívá standard URN:NBN a je řízen Národní knihovnou České republiky (NK ČR), která systém spravuje a koordinuje ve spolupráci s dalšími institucemi. Služby systému ČIDLO jsou dostupné na adrese <https://resolver.nkp.cz>. Služba ČIDLO je založena na softwaru CZIDLO. S uživateli automatizovaně komunikuje pomocí aplikačního rozhraní (Application Programming Interface – API). Jedná se o **RESTful webové API**. Pro autentizaci je použita metoda *Basic access authentication*. Operace, které jsou autorizované, jsou dostupné jen přes *HTTPS*, čímž je zabráněno zachycení loginu a hesla, které jinak metoda posílá v otevřené formě.

Digitální dokument

Digitální (či přesněji vymezeno v rámci této metodiky digitalizovaný) dokument je komplexní pojem, který má v rámci metodiky a informační vědy dva významy. Primárně bude pojem používán ve smyslu obecného označení digitálního objektu vzniklého digitalizací (mimo tuto metodiku patří do definice i objekty vzniklé přímo v digitální podobě, tzv. digital born dokumenty nebo e-borny) nebo-li množiny počítačových souborů v určitém formátu, které jako celek, spolu se strukturálními metadaty, tvoří, při zobrazení v příslušné aplikaci (např. internetový prohlížeč), obraz celé tištěné předlohy včetně zachování posloupnosti stran. V rámci systému ČIDLO je digitální dokument definován jako *statický digitální dokument*. Dokumentu, jehož obsah se nadále nemění, lze přidělit trvalý identifikátor URN:NBN. Digitální dokument je to, co je předmětem identifikace. Digitální dokument digitálně reprezentuje jednu intelektuální entitu a je popsán týmiž technickými a bibliografickými metadaty (tato dědí po intelektuální entitě). Digitální dokument má zpravidla dvě verze – archivní (pro archivaci v digitálním repozitáři) a uživatelskou (pro zpřístupňování v digitální knihovně), obě verze jsou považovány z hlediska autenticity za shodné a mají tentýž identifikátor URN:NBN. V přesně vymezených případech bude pojem používán jako přesněji vymezený, jako označení abstraktního modelu digitálního dokumentu skládajícího se ze tří vrstev: intelektuální entity, digitální reprezentace a digitální instance, přičemž pojem digitální reprezentace odpovídá popsanému obecnému vnímání digitálního dokumentu. V rámci systému ČIDLO odpovídá pojem digitální dokument digitální reprezentaci.

Digitální instance

Umístění uživatelské kopie dokumentu v digitální knihovně, ze kterého může uživatel získat přístup k celému uživatelskému dokumentu. Digitální instance je identifikována pomocí URL

adresy. Digitální instance je lokalizována na „splash page“, tedy na úvodní webové stránce k digitálnímu dokumentu, obsahuje odkazy na jednotlivé soubory tvořící digitální reprezentaci a jsou na ní uvedena metadata k digitálnímu dokumentu. Digitálních instancí jedné reprezentace může být více, pokud jeden vlastník zpřístupní digitální dokument skrze více digitálních knihoven.

Digitální knihovna

Systém, ve kterém jsou uživatelské kopie dokumentů zpřístupňovány uživatelům.

Digitální repozitář

(synonymum: digitální úložiště)

Systém pro dlouhodobé uchovávání archivních kopií digitálních dokumentů. Tento systém musí splňovat požadavky archivu OAIS.

Digitální reprezentace

Digitální reprezentace je množina počítačových souborů, která vznikla digitalizací a která reprezentuje intelektuální entitu. Pokud z jedné intelektuální entity vytvoří (zdigitalizuje) více institucí digitální reprezentaci, pak se jedná o různé digitální reprezentace a každá získá jiný identifikátor. Podmínkou získání identifikátoru URN:NBN je tedy vlastnictví digitální reprezentace. Pokud je digitální reprezentace dostupná ve více formátech, stále se jedná o jednu digitální reprezentaci a všechny verze mají jeden společný URN:NBN identifikátor a jsou lokalizovány jako jedna digitální instance. K digitální reprezentaci se přiřazuje URN:NBN identifikátor. Digitální reprezentace terminologicky odpovídá digitálnímu dokumentu.

DIP

Výstupní informační balíček (dissemination information package)

Informační balíček odvozený z jednoho nebo více balíčků AIP a zaslaný archivem OAIS koncovému uživateli jako odpověď na jeho požadavek vůči tomuto archivu

FRBR

Koncepční model Functional Requirements for Bibliographic Records – Funkční požadavky na bibliografické záznamy.

Identifikátor

Znakový řetězec, který v rámci určitého kontextu jednoznačně označuje nějaký objekt.

Identifikátorová metadata

Metadata, která jsou uložena v databázi resolveru spolu s identifikátorem URN:NBN a která popisují původní archivní dokument.

Intelektuální entita (IE)

Intelektuální entita je ucelenou množinou informací z hlediska správy i popisu. Intelektuální entitou může být pro potřeby trvalé identifikace kniha, článek v periodiku, periodikum (seriál) jako číslo nebo jednotlivý ročník, kvalifikační práce, stať apod. K intelektuální entitě se přiřazují bibliografická metadata.

Nástupnický archivní dokument

Archivní dokument, který je potomkem původního archivního dokumentu první nebo vyšší generace. Nástupnický archivní dokument vzniká v digitálním repozitáři formátovou migrací předešlého archivního dokumentu a je obsažen v nové verzi balíčku AIP.

Nestandardní dokument

Digitální dokument, který nebyl vytvořen podle standardů digitalizace NDK. Může se jednat o dokument, který vznikl předtím, než vstoupily v platnosti standardy digitalizace NDK, nebo který vznikl v době platnosti standardů digitalizace NDK, ale podle jiných standardů. I tento dokument může za určitých okolností obdržet perzistentní identifikátor URN:NBN, ale funkčnost bude v rámci systému omezena. Především bude omezena schopnost resolvovat všechny dostupné digitální instance dokumentu, zpravidla bude odkazována jen kopie v digitální knihovně přímého vlastníka nebo pověřeného archivátora.

OAIS

otevřený archivační informační systém (open archival information system)

Archiv, který tvoří uskupení lidí a systémů, jež přijalo odpovědnost za uchovávání informací a jejich zpřístupňování určené skupině, přičemž toto uskupení může být součástí většího celku.

Původní archivní dokument

Archivní dokument, který byl uložen do digitálního repozitáře jako první verze balíčku AIP.

Původní uživatelský dokument

Uživatelský dokument ve formátu, ve kterém byl vytvořen při produkci, je tvořen uživatelskými kopiemi.

Registrátor

Instituce, která buď sama, nebo prostřednictvím zhotovitele získává identifikátory URN:NBN a zapisuje je do metadat vázajících se k archivním i uživatelským verzím dokumentů.

SIP

Submission information package

Informační balíček, který dodává tvůrce do archivu OAIS tak, aby mohl být využit při sestavení nebo aktualizaci jednoho nebo více AIP a/nebo přidružených popisných informací.

Standardy digitalizace NDK

Standardy digitalizace, která zavedla NK ČR pro projekt Národní digitální knihovny. V současnosti jsou tvořeny DMF (definice metadatových formátů), standardy pro obrazová data, pravidly popisu a jsou určeny pro digitalizaci monografií a periodik. Jsou dostupné na webu Národní digitální knihovny: <http://www.ndk.cz/standardy-digitalizace>

Technická metadata

Metadata popisující technické vlastnosti obrazových souborů ve standardu MIX.

Uživatelský dokument

Uživatelská kopie digitálního dokumentu zpřístupňovaná uživatelům, obvykle v nižší obrazové kvalitě než archivní dokument.

5. IDENTIFIKÁTORY

5.1 Přehled stavu - identifikátory v praxi knihoven

5.1.1 Identifikátory tištěných dokumentů

Identifikátory obecně slouží k jednoznačnému označení objektu v určitém prostředí. V prostředí knihoven a produkce a distribuce tištěných knih a periodik mají relativně dlouhou tradici identifikátory tištěných dokumentů, zejména identifikátory ISBN a ISSN. ISBN je tzv. perzistentní identifikátor, tj. takový, který je trvale užíván k identifikaci právě té publikace, které byl původně přidělen (tj. nesmí být přidělen znovu, a proto musí být trvale udržován v centrální databázi). Globální trvalé identifikátory mají standardizovanou syntax a jsou přidělovány na základě určitých pravidel, mezi která obvykle patří definice toho, čemu může být identifikátor přidělen. Například identifikátor ISBN může být přidělen pouze knize na úrovni vydání, nikoliv exempláři³, tj. ISBN identifikuje dokument z hlediska jeho intelektuálního obsahu (a sekundárně též z hlediska specifických znaků vydání).

Identifikátor ISBN slouží knihovnám zejména pro potřeby správy, popisu a uchovávání tištěných knih dokumentů nebo sdílení bibliografických záznamů; vydavatelům, distributorům a prodejcům knih zase poskytuje prostředek pro propagaci, distribuci nebo nabízení knih čtenářům. Uživatelům knihoven umožňuje identifikátor ISBN rychlé a přesné vyhledání záznamu požadované knihy v katalogu. Kupujícím umožňuje jednoznačnou identifikaci produktu, který si koupili nebo chtějí zakoupit či objednat.

Z důvodů globalizace (nejen knižního) trhu i mezinárodní spolupráce knihoven ve 20. století a požadavků dlouhodobé využitelnosti identifikátorů jsou identifikátory ISBN a ISSN globální (jednoznačné v mezinárodním měřítku) a trvalé (identifikátor jednou přidělený určitému dokumentu nesmí být nikdy přidělen jinému). Globálnost, perzistence a standardizace (pravidla) identifikátoru zaručují jeho dlouhodobou užitečnost při správě a vyhledávání dokumentů.

Úspěch obou zmíněných identifikátorů spočívá v tom, že jsou řízeny dobře fungujícími identifikačními systémy. Identifikátor sám, bez identifikačního systému, nemůže plnit potřebné funkce v širší komunitě uživatelů. Identifikátory ISBN a ISSN jsou přidělovány centralizovanými hierarchizovanými systémy (mezinárodní agentura a síť podřízených

³ Všechny exempláře téhož vydání mají totéž ISBN.

národních agentur). Tyto systémy vymezují pravidla užití identifikátorů a řídí a kontrolují jejich přidělování.

Pravidla užití zahrnují definici datového modelu, tedy toho, co je předmětem identifikace. Například identifikátor ISBN je přidělován každému vydání (prvnímu, druhému atd.) knihy na úrovni titulu, tzn. všem exemplářům vydání knihy, které mají tentýž obsah. V rámci vydání se dále odlišuje jiná forma (vázané vydání, brožované vydání) a vydavatel. Každá jiná forma téhož vydání dostane jiný identifikátor ISBN. Pokud odlišní vydavatelé vydají tutéž knihu, dostane každá z nich jiné ISBN. Pravidla užití mj. vymezují, že identifikátor ISBN musí být vytištěn v knize samotné, je tedy trvale spojen s touto knihou, což usnadňuje identifikaci konkrétní knihy.

Tento model byl navržen tak, aby vyhovoval potřebám těch, kteří budou identifikátory ISBN využívat. Různá vydání mohou obsahovat odlišné verze díla (např. přepracované vydání) nebo jinou typografickou úpravu. Totéž vydání může existovat v různých formách (např. vázané nebo brožované vydání). Model systému ISBN umožňuje zachytit tyto rozdíly, protože jeho původní funkcí je identifikace specifického sériového produktu, dnes nezbytná potřeba z hlediska distribuce a prodeje knih (to platí obecně o jakémkoliv hromadně prodávaném zboží - viz čárový kód zboží). Naproti tomu jednoznačnou identifikaci jednotlivého exempláře vydání knihy ISBN neumožňuje (totéž platí o čárovém kódu zboží), neboť takový způsob identifikace není potřebný pro prodej a distribuci.

Ačkoliv motivace vzniku systému ISBN byla komerční, existence tohoto identifikátoru je výhodná i pro knihovny, například pro potřeby katalogizace nebo sdílení záznamů. Z tohoto důvodu jsou také některé národní agentury provozovány tamními národními knihovnami.

V knihovnách se současně s ISBN užívá identifikátor exempláře, tj. signatura. Signatura slouží k určení lokace knihy – podle signatury většinou nalezneme knihu ve volném výběru nebo v depozitáři. Signatura není globální identifikátor (signatura je jedinečná jen v rámci dané knihovny), ani perzistentní identifikátor (knihy může dostat novou signaturu), ani nepodléhá rozšířené standardizaci (každá knihovna si může zvolit vlastní syntax pro signatury).

5.1.2 Identifikátory a digitální dokumenty

1

V případě digitálních dokumentů se možnosti využití identifikátorů rozšiřují⁴. Digitální dokumenty jsou často online, tj. přístupné prostřednictvím internetové sítě. Pro online dokumenty platí, že na základě identifikátoru digitálního dokumentu je možné vyhledat nejen bibliografické údaje popisující tento dokument, ale také k němu získat okamžitý přístup, ať již jeho prohlížením na dané webové stránce nebo stažením z ní do počítače uživatele. Nejčastější způsob identifikace digitálních dokumentů zpřístupňovaných v internetové síti je URL adresa webové stránky.

V současné době většina uživatelů vyžaduje, aby byly digitální dokumenty přístupné na internetu, a převážná většina producentů digitálních dokumentů jim vychází vstříc. Požadavek, aby byl digitální dokument přístupný internetovou sítí, je současně požadavkem po tom, aby identifikátor digitálního dokumentu plnil funkci, kterou identifikátory tištěných dokumentů nikdy plnit nemohly - zprostředkoval okamžitý přístup k celému dokumentu.

Způsob identifikace digitálního dokumentu prostřednictvím URL adresy však není stabilní, neboť dokumenty mohou být (a ve většině případů jsou) přesouvány na jinou webovou stránku či jinam v rámci dané stránky.

Způsob identifikace dokumentu prostřednictvím URL adresy je de facto identifikací umístění dokumentu, nikoliv dokumentu samotného. Dokument je na dané internetové adrese dostupný z kteréhokoliv počítače na světě (s platnými legislativními, případně komerčními omezeními), a z tohoto důvodu se také stalo URL nejužívanějším digitálním identifikátorem současnosti. Zjednodušeně lze říci, že URL je jakási „globální signatura“, protože URL je identifikátor lokace. Tato „globální signatura“ je sice na rozdíl od signatur knih vysoce standardizovaná, nicméně není perzistentní – stejně jako knihovní signatura se může změnit i internetová lokace dokumentu.

Pro trvalou využitelnost dokumentů v internetové síti je nutná existence systému trvalé identifikace, který je založen na globálním a perzistentním identifikátoru vlastního dokumentu (nikoliv jeho lokace). Svět digitálních dokumentů pracuje s celou řadou identifikátorů. Pouze některé jsou perzistentní, tedy trvalé, poskytující jednoznačnou, dlouhodobou identifikaci. Z výše naznačených důvodů jsou již od poloviny 90. let 20. století navrhovány a zaváděny

⁴ Jak bylo řečeno v části definující pojmy používané metodikou, jako digitální dokument je zde označován digitální objekt vzniklý digitalizací, který v abstraktním modelu digitálního dokumentu odpovídá rovněž digitální reprezentaci. V rámci dokumentace systému ČIDLO je digitální reprezentace označována jako digitální dokument.

nové systémy identifikace digitálních dokumentů, které se pokouší řešit tento problém a které nejsou založeny primárně na URL adrese. Jeden z těchto systémů, systém URN:NBN (Universal Resource Name: National Bibliography Number), byl vybrán pro identifikační systém ČIDLO, který je českou národní implementací systému URN:NBN.

Mezi další rozšířené identifikační systémy patří Handle, DOI (Digital Object Identifier), PURL (Persistent Uniform Resource Locator), ARK (Archival Resource Key).

Předem je třeba říci, že není nutné, aby všechny instituce využívaly stejný typ identifikátoru, protože se jejich charakter např. nehodí pro odlišné fondy. A na druhou stranu ničemu nebrání, pokud bude jeden dokument označen více identifikátory, pokud se tak bude dít vždy ve vymezeném kontextu.

Široké mezinárodní užití má systém Handle (www.handle.net), který vyvinula a provozuje CNRI (Corporation for National Research Initiatives) jako součást obecné architektury pro jednoznačnou identifikaci digitálních objektů. Služeb tohoto systému může využívat jakákoliv instituce, která může využívat i další softwarové nástroje CNRI včetně databáze spravující poskytnuté identifikátory. Samotné využívání systému Handle je zpoplatněno, zatím mírnou taxou. Nevýhodou systému je především nemožnost na základě identifikátoru ihned identifikovat vlastníka, případně geografickou příslušnost.

Velmi rozšířený, ale zároveň omezený, je systém DOI (<http://www.doi.org/>) založený na bázi Handle, který však poskytuje mnohem více funkcí a dokáže identifikovat různé typy digitálních objektů (obchodní smlouvy, patenty atd.). Systém vznikl pro identifikaci digitálních objektů pro potřeby komerčních vydavatelů. Syntax DOI byla specifikována normou ANSI/NISO Z39.84-2000. Identifikátor DOI je využíván především vydavateli vědeckých on-line publikací, přičemž dnes lze říci, že velká část vědeckých článků je identifikována právě tímto systémem (DOI identifikuje primárně samotný článek a ne titul). Pro získání identifikátoru je třeba odevzdávat poměrně značné množství metadat, řídit se komplexními pravidly (povinnost registrovat všechny přidělené identifikátory u agentury, která následně zajišťuje i směrování) a především, což vyplývá z faktu, že DOI je určen především pro podporu komerčního vydávání, je zpoplatněn za přidělení každého jednotlivého identifikátoru. Toto zpoplatnění, jakož i komplexnost procesu přidělování, vylučuje systém DOI z provozu běžných knihoven zabývajících se masovou digitalizací.

Identifikátor PURL (<http://purl.org>) byl vyvinut Online Computer Library Center. Je založen na klasické URL, ale s pomocí softwaru resolveru je schopen automaticky přesměrovávat na aktuální umístění dokumentu. Funkčně celý proces probíhá tak, že

dokument dostane URL odkazující na službu PURL a teprve z ní je uživatel odkázán na aktuální umístění. Instituce mohou mít i vlastní resolver.

Mezi užívané patří i identifikátor ARK, jenž je opět postaven na bázi URL. ARK umožňuje identifikovat různé typy nejen digitálních objektů. Může být využit pro označení různých intelektuálních entit, knih, ale i softwarů, databází apod. Jeho schéma je poměrně komplikované, vedle povinné obsahuje i volitelnou část umožňující směrování.

2 Dlouhodobé uchovávání a trvalé identifikátory

Z hlediska dlouhodobého uchovávání digitálních dokumentů se objevily nové výzvy, které ve světě tištěných dokumentů neexistovaly. Forma tištěného dokumentu je tvořena materiálním objektem a uchovávání tištěné knihy nebo periodika znamená opatření, jak zachovat tento materiální objekt, mezi něž patří omezení přístupu k němu a zpomalení jeho degradace (která je nevyhnutelným fyzickým procesem). Digitální dokumenty jsou tvořeny bity, které jsou uloženy na datovém nosiči. Datové nosiče jsou sice také materiální objekty (např. pevný disk), a tedy také degradují, ale bity uloženy na těchto nosičích jsou na těchto nosičích nezávislé, a je tak možné kopírovat je na nové datové nosiče v případě, že původnímu datovému nosiči hrozí degradace nebo zastarání (konec podpory počítačovými systémy). Kopírování bitů na nové nosiče je základním opatřením dlouhodobého uchovávání digitálních dokumentů. Bity jsou zpravidla uloženy ve formě počítačových souborů, které jsou v určitém datovém formátu. Tyto datové formáty jsou ohroženy zejména zastaráním - zastaralost formátu znamená, že již neexistují programy, které by dokázaly soubory v těchto formátech otevřít a adekvátně zobrazit. Nyní je již obecně přijatou skutečností, že digitální dokumenty nebude možné v budoucnosti dále uchovávat ve stejných datových formátech, ve kterých byly původně vytvořeny nebo uloženy do digitálního úložiště. Předpokládá se, že formáty digitálních dokumentů bude nutné převádět do nových formátů, které umožní, aby digitální dokument bylo v novém počítačovém prostředí možno adekvátně zobrazit. V některých případech je tento požadavek aktuální již nyní. Některé knihovny, které v 90. letech digitalizovaly své tištěné dokumenty do formátu TIFF, zvažují převod do formátu JPEG2000, některé knihovny jej dokonce již uskutečňují. (Palmer, 2013)

Převod do jiného formátu by měl z hlediska požadavků dlouhodobého uchovávání současně zachovat obsah dokumentu vč. zachování podstatných vlastností daného digitálního dokumentu.

Z hlediska identifikace tato skutečnost nastoluje následující problémy. Pokud se změní formát digitálního dokumentu a nová verze digitálního dokumentu dostane nový identifikátor, nebude moci uživatel na základě identifikátoru, přiděleného původnímu dokumentu, vyhledat novou verzi tohoto dokumentu. V případě ISBN, jak jsme viděli, je jedním z prvků datového modelu forma vydání. Datový model systému ISBN změnu formy knihy neřeší. Původní pravidla systému ISBN se vztahovala pouze na tištěné knihy a změna formy vydané tištěné knihy není možná - kniha zůstává stejným fyzickým objektem. Poté, co začali vydavatelé vydávat kromě tištěných knih také online publikace, systém ISBN rozšířil možnosti svého užití i na tento nový typ knihy. V případě, že jedno vydání je vydáno v různých formátech (např. epub nebo PDF), každý formát musí dostat nový identifikátor ISBN. To platí i v případě odlišných verzí téhož formátu (e-book ve formátu PDF s ochranou DRM dostane jiné ISBN než e-book v PDF bez ochrany DRM). Knihovny, které digitalizují své tištěné dokumenty, se do určité míry podobají klasickým vydavatelům - do té míry, že jsou producenty i vlastníky digitalizovaných dokumentů. Nemají však zpravidla, na rozdíl od klasických vydavatelů, práva copyrightu k tištěným předlohám, ani licenci na autorská práva k těmto předlohám. Cílem produkce navíc nebývá prodej digitalizovaných dokumentů, ale zejména lepší služby pro uživatele a ochrana tištěných dokumentů (např. omezení přístupu k tištěnému dokumentu tím, že je nabízena jeho digitální kopie). Z hlediska uživatelů je pak důležité, aby byl dokument zpřístupnitelný i v budoucnosti, což pravděpodobně bude znamenat změnu jeho formátu.

3 Digitalizace a trvalé identifikátory

V digitalizační praxi velké části knihoven převládá přístup, ve kterém má výsledek digitalizace dvě podoby - archivní kopie a uživatelské kopie. Archivní a uživatelské kopie se zpravidla liší formátem nebo jeho verzí (např. JP2 v bezztrátové nebo ztrátové kompresi), nicméně jejich obsah zůstává stejný - je jím obraz tištěné předlohy. Vedle tohoto přístupu existuje druhé řešení, a sice generovat uživatelské kopie z archivních, většinou na vyžádání uživatele. Tato metodika popisuje řešení pro převládající postup, ale pokud by někdo zvolil minoritní postup, platí pro něj z hlediska perzistentních identifikátorů v zásadě stejná pravidla. Perzistentní identifikátor URN:NBN musí být při generování uživatelské kopie přebrán do metadat uživatelského dokumentu.

U uživatelských kopií vyvstává tentýž problém, jako v případě dokumentu v původním formátu a následnickém dokumentu v jiném formátu, a sice zda tyto odlišné kopie

označit stejným, nebo odlišným identifikátorem, a pokud stejným, pak jakým způsobem přesně definovat předmět identifikace.

Důvodem současné produkce archivní a uživatelských kopií je to, že jejich formáty jsou vybírány s ohledem na užití dokumentu. Formát archivní kopie je vybírán na základě vhodnosti pro dlouhodobé uchování. V odborné komunitě se za takový formát považuje formát, který splňuje řadu kritérií, například není proprietární (tj. jeho užití není omezeno právy duševního vlastnictví), je dobře zdokumentovaný a standardizovaný a bude z něj v budoucnosti snadnější provádět převody do jiných formátů.

Naproti tomu formát uživatelské kopie se volí z hlediska potřeb uživatelů a možností digitální knihovny. Například soubory ve ztrátové kompresi jsou obvykle menší velikosti, a proto je uživatel může rychleji načíst, a podobně méně zatěžuje provoz digitální knihovny a její úložné kapacity. Formáty pro uživatele jsou také vybírány z hlediska toho, že jsou podporovány v běžně užívaných aplikacích (např. internetové prohlížeče) nebo na oblíbených zařízeních (např. tablety).

5.2 Identifikátor URN:NBN

Identifikátor URN:NBN může být tvořen pouze alfanumerickými znaky anglické abecedy (0-9; a-z), dvojtečkou a spojovníkem. Pro systém ČIDLO je využita varianta se systémem podřízených jmenných prostorů (uvedená ve standardu URN:NBN). Syntax identifikátoru URN:NBN v systému ČIDLO lze schematicky vyjádřit jako:

A:B-C

Vždy platí, že:

A = urn:nbn:cz (urn.nbn je obecný předpis; cz kód jmenného prostoru)

B = kód registrátora (1-6 alfanumerických znaků).

V případě knihovny se záznamem v ADR platí, že B = sigla (šestimístný alfanumerický znakový řetězec), kromě několika výjimek (Národní knihovna ČR = nk [pro digitalizaci od roku 2012], Moravská zemská knihovna = mzk, Vědecká knihovna v Olomouci = vkol).

Kód registrátorů (část B) bez záznamu v ADR může obsahovat 1-5 alfanumerických znaků, jejich výběr je nutno konzultovat s kurátorem resolveru, který kód následně schvaluje a eviduje v databázi. Výsledný identifikátor v rámci této varianty tak může být tvořen 19-23

znaky.

C = 6 alfanumerických znaků

Metodika doporučuje zapisovat identifikátor malými písmeny. Z důvodu snadnější čitelnosti lidským uživatelem. (Nulu „0“ a písmeno „O“ lze v některých znakových sadách snadno zaměnit).

Výsledný identifikátor má pak vždy toto schéma a skládá se z 19 až 24 znaků (příklad pro Národní lékařskou knihovnu):

A						:	B						-	C									
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
u	r	n	:	n	b	n	:	c	z	:	a	b	a	0	0	8	-	0	2	b	4	5	a

Obrázek 1: Příklad identifikátoru pro NLK; urn:nbn:cz:aba008-02b45a

Příklad pro Národní knihovnu ČR

A						:	B	-	C										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
u	r	n	:	n	b	n	:	c	z	:	n	k	-	0	0	1	4	t	a

Obrázek 2: Příklad identifikátoru pro NK

V ČR zajišťuje přidělování a správu identifikátorů URN:NBN systém ČIDLO provozovaný Národní knihovnou ČR.

5.3 Východiska systému ČIDLO

Systém ČIDLO je postaven na následujících předpokladech:

Knihovny, které digitalizují své tištěné předlohy, mají obvykle za cíl vytvořit odpovídající obraz tištěné předlohy, poskytnout jej uživatelům a zachovat jej pro budoucí uživatele.

Obraz fyzické předlohy je obsahem digitalizovaného dokumentu, který není závislý na konkrétním grafickém formátu. Knihovny, které digitalizují své tištěné fondy, aby je následně dlouhodobě uchovávaly, nejsou v pozici klasických vydavatelů, kteří vydávají knihy v různých formátech z důvodu vytvoření širší nabídky produktů pro čtenáře. Volba formátů pro archivní a uživatelské dokumenty (a do budoucna pro nástupnické archivní dokumenty a nové verze uživatelských dokumentů) je motivována primárně technickými omezeními, nikoliv komerčním záměrem.

Předmětem identifikace v systému ČIDLO z těchto důvodů proto není obsah a forma (formát) digitalizovaného dokumentu, ale pouze obsah, tj. obraz fyzické předlohy. Fyzickou předlohou se rozumí předloha na úrovni provedení podle FRBR. To znamená, že na obrazové odlišnosti mezi jednotlivými exempláři fyzické předlohy se nehledí (tyto odlišnosti mohou být způsobeny například různým fyzickým stavem různých exemplářů nebo dodatečně přidanými prvky, jako je například razítko knihovny).

Tato volba zohledňuje v první řadě hledisko uživatele dokumentu. Uživatel má v tomto případě možnost získat přístup k témuž obsahu, bez ohledu na to, zda jde o původní archivní dokument nebo nástupnický archivní dokument, nebo různé verze uživatelského dokumentu. Taktéž uživatele nezajímají odlišnosti jednotlivých exemplářů, které jsou výsledkem degradace nebo zacházení s exemplářem v knihovně od doby jeho vytištění..

Tato volba je též výhodná pro správu dokumentů, protože umožňuje na základě identifikátoru URN:NBN propojit archivní a uživatelský dokument s tímž obsahem, nebo původní archivní dokument a nástupnické archivní dokumenty v digitálním repozitáři. Pro identifikaci různých verzí dokumentu lze využít jiné identifikátory (např. UUID). Tyto jiné identifikátory uživatele nezajímají, protože neidentifikují obsah. Nicméně tyto jiné identifikátory lze ve spojení s identifikátorem URN:NBN využít ke koordinaci správy dokumentů s tímž obsahem, které jsou uloženy v různých systémech.

Stejnou tištěnou předlohu může zdigitalizovat více institucí. Ačkoliv obraz bude stejný, jeho vlastník bude odlišný. Pro uživatele znamená jiný vlastník jiná přístupová omezení a informaci o odpovědnosti za kvalitu digitalizace. Totéž platí v případě, že stejný vlastník zdigitalizuje tutéž entitu v rámci jiných projektů. V takovém případě sice bude obraz i vlastník stejný, ale digitalizace může probíhat podle jiných pravidel (např. podle budoucích

standardů digitalizace NDK), což může mít opět vliv na kvalitu obrazu. **Je tedy užitečné tyto odlišnosti zaznamenat užitím jiného identifikátoru URN:NBN.** Z praktického hlediska by navíc mohlo být velmi obtížné využít pro identifikaci již přidělený identifikátor - vyžadovalo by dohledávání původního identifikátoru, což by mohlo značně zdržovat produkci.

Volba přidělit identifikátor již při produkci digitalizovaného dokumentu je nevyhnutelná - po produkci se různé verze téhož dokumentu odesílají do různých systémů a pozdější identifikace se tím stává prakticky vyloučenou.

Identifikovaný dokument musí splňovat tyto požadavky:

- a) trvalá hodnota a význam dokumentu z hlediska digitálního kulturního dědictví, jde primárně o dokumenty textové povahy (po konzultaci s kurátorem resolveru),
- b) ucelenost dokumentu (smysluplný a samostatný informační celek)
 - a. základní typy ucelených dokumentů: monografie, články / čísla / ročníky seriálu (ne tituly seriálů), mapy, hudebniny, vysokoškolské závěrečné práce, případně další typy pouze po konzultaci s kurátorem resolveru,
- c) existence samostatných bibliografických a technických metadat vztahujících se právě k tomuto dokumentu,
- d) přiměřenost rozsahu dokumentu (např. nikoliv celá sbírka dokumentů, nikoliv jedna stránka vícestránkové dokumentu apod.),
- e) neměnnost obsahu dokumentu (statický dokument).

6. SYSTÉM ČIDLO

ČIDLO je systém pro trvalou identifikaci digitálních dokumentů českého kulturního dědictví, který využívá standard URN:NBN a je řízen Národní knihovnou ČR, která tento systém spravuje a koordinuje ve spolupráci s dalšími institucemi, přičemž tato spolupráce je nezbytnou podmínkou úspěšného fungování celého systému. ČIDLO slouží jednak knihovnám a dalším institucím z oblasti českého kulturního dědictví, a to pro jejich potřeby trvalé identifikace jejich digitálních dokumentů (procesy dlouhodobé správy, archivace a zpřístupňování digitálních dokumentů), jednak uživatelům, a to jako prostředek pro zajištění trvalého zpřístupňování digitálních dokumentů v internetové síti (řešení problému nestability a proměnlivosti URL adres tím, že resolver nabízí přesměrovací služby) nebo jako prostředek pro zajištění důvěryhodnosti citační praxe (řešení problému ověřování autenticity citovaných dokumentů).

ČIDLO se skládá z pravidel, technických podsystémů (zejména resolver, API, OAI-PMH data provider), zapojených institucí a jejich pověřených zaměstnanců, identifikátorů URN:NBN pro český jmenný prostor (tj. identifikátorů začínajících řetězcem „urn:nbn:cz“) a k nim přidružených metadat. Roli centrální autority vykonává – v souladu se standardem URN:NBN – výlučně Národní knihovna ČR jakožto centrální autorita pro Českou republiku. Jádrem fungování systému ČIDLO je databáze obsahující identifikátory URN:NBN a identifikátorová metadata a resolver zajišťující přesměrování na aktuální umístění identifikovaného dokumentu.

Problémem, který systémy trvalé identifikace řeší, je také skutečnost, že digitální dokumenty často přestanou být dostupné online. Pak již je obtížné ověřit, zda daný dokument někdy existoval, což může narušovat důvěryhodnost citační praxe. Právě v podpoře důvěryhodnosti citací digitálních dokumentů spočívá jedna z funkcí systému ČIDLO. I v případě smazání identifikovaného dokumentu z digitální knihovny, což by se teoreticky nemělo nikdy stát, ale stává se, zůstane záznam v systému a vždy tak budou dohledatelné informace o dokumentu (identifikátorová metadata), který byl citován.

6.1 Resolver

Resolver je aplikace, která zajišťuje

- a) přidělování a správu jednoznačných unikátních perzistentních identifikátorů založených na standardu URN:NBN v českém prostředí,

- b) trvalé udržování databáze identifikátorů, identifikátorových metadat a URL adres (metadata udržuje i v případě zániku nebo on-line nedostupnosti vlastního identifikovaného dokumentu)
- c) poskytování a registrování metadat vázaných k identifikátorům,
- d) přesměrovávání na aktuální URL adresy digitálních dokumentů do digitálních knihoven odpovědných za zpřístupňování daných dokumentů,
- e) sklizení aktuálních URL adres z daných digitálních knihoven ve spolupráci s těmito knihovnami,
- f) komunikační rozhraní pro přidělování identifikátorů a získávání metadat k nim vázaných od registrátora (REST pro automatizované přidělování, webové rozhraní pro manuální přidělování) a sklizení aktuálních URL adres dokumentů (OAI-PMH) v součinnosti s danou digitální knihovnou,
- g) správu uživatelských účtů (zpravidla jeden uživatel z dané instituce spravuje účet registrátora).

7. PODMÍNKY ÚČASTI V SYSTÉMU ČIDLO

Registrátorem s právem spravovat svůj jmenný prostor a přidělovat svým digitálním dokumentům identifikátory URN:NBN se může stát každá knihovna registrovaná v ADR a dále každá další instituce spravující knižní fond nebo fond podobné povahy (kartografické dokumenty, hudebniny, šedá literatura). Registrátorem tedy může být každá instituce s patřičnými kompetencemi v oblasti správy a zpřístupňování digitálních dokumentů v rozsahu popsaných fondů. V případě, že je instituce digitalizující své fyzické předlohy vázána povinností ukládat své digitalizované dokumenty do LTP NK ČR, nebo má záměr je zde ukládat, pak pro ni platí povinnost řídit se standardy digitalizace NDK.

Od verze standardů digitalizace NDK (verze 1.5 pro periodika a 1.1 pro monografie) je zavedena povinnost nechat si přidělit identifikátor URN:NBN, což znamená, že se automaticky stává způsobilou pro účast v systému ČIDLO. V tomto případě je instituce povinna dodržovat pravidla představená v této metodice.

Každá další instituce se může stát registrátorem (i v případě, že není knihovnou) pokud ji nevyklučují následující omezení.

Skutečnosti, které vylučují možnost stát se registrátorem:

- žadatel není vlastníkem digitalizovaného dokumentu,
- žadatel není instituce z oblasti správy kulturního nebo vědeckého dědictví nebo podobné oblasti,
- fyzická předloha digitalizovaného dokumentu není bohemikum (pro potřeby systému ČIDLO je bohemikum vymezeno jako takové knižní dílo nebo dílo ve správě knihoven, které má souvislost s českou kulturou a územím a to i kdyby se jednalo o souvislost vzniklou místem uložení; v případech pochybností rozhoduje kurátor),
- digitalizovaný dokument nemá trvalou hodnotu a význam z hlediska českého národního dědictví (v případě nejasností ohledně určení tohoto kritéria rozhodují stanoviska kurátora služby ČIDLO),
- obsah digitalizovaného dokumentu není určen k zpřístupňování široké veřejnosti (tajné materiály, průmyslová dokumentace atd.),
- žadatel nemá záměr digitalizovaný dokument dlouhodobě uchovávat,
- žadatel nemá záměr digitalizovaný dokument zpřístupňovat uživatelům na nekomerční bázi prostřednictvím internetové sítě.

Žadatel, kterého se výše uvedené skutečnosti netýkají, se může stát registrátorem, **pokud se zaváže plnit následující podmínky:**

- bude vytvářet své digitalizované dokumenty na základě aktuální verze standardů digitalizace NDK nebo bude převádět své nestandardní dokumenty do podoby odpovídající schválené verzi standardů digitalizace NDK,
- v případě, že své digitalizované dokumenty nebude vytvářet podle standardů digitalizace NDK, je zapojení i tak možné, pokud bude schopen odevzdávat minimálně požadovaný rozsah metadat. (Tato metodika primárně předpokládá, že digitalizované dokumenty budou vznikat podle standardů digitalizace NDK a pro tyto je také nastavena),
- bude dodržovat postupy stanovené touto metodikou,
- zajistí si technické a další prostředky pro komunikaci s resolverem,
- bude své dokumenty zpřístupňovat pomocí digitální knihovny
- digitální knihovna, ve které budou zpřístupňovány digitalizované dokumenty, bude disponovat protokolem OAI-PMH pro sklízení metadat,
- zajistí uchování dokumentů, musí se zaručit, že identifikovaný dokument nesmaže.

Podmínka vytvářet digitalizované dokumenty podle standardů digitalizace NDK znamená, že systém ČIDLO je určen primárně pro identifikaci nově vznikajících digitalizovaných dokumentů. Současně však je možné identifikovat nestandardní dokumenty, a to v procesu jejich převodu do podoby podle standardů digitalizace NDK nebo pokud jejich podoba bude odpovídat minimálním požadavkům této metodiky a dokumenty budou splňovat nároky na dlouhodobé uložení.

8. MODEL ROLÍ A ODPOVĚDNOSTÍ V SYSTÉMU ČIDLO

Účastníky systému ČIDLO jsou registrátoři a centrální autorita. Centrální autorita je správcem systému ČIDLO a přiděluje identifikátory URN:NBN. Registrátor je instituce, která je vlastníkem digitalizovaného dokumentu. Centrální autorita přiděluje registrátorovi identifikátory URN:NBN. Centrální autoritu představuje Národní knihovna ČR. Jejím zástupcem je kurátor systému ČIDLO (dále jen kurátor). Registrátor může být zastupován digitalizační firmou či jinou institucí, pokud ji registrátor zmocní.

Identifikátory URN:NBN jsou užívány v následujících systémech: resolver, digitalizační systém (systém pro produkci digitalizovaných dokumentů), katalog, digitální repozitář, digitální knihovna.

Resolver je online aplikace, kterou provozuje centrální autorita. Hlavní úlohou resolveru je přidělování a správa identifikátorů URN:NBN a přesměrovávání z identifikátoru URN:NBN na aktuální URL adresu do digitální knihovny, ve které je identifikovaný dokument zpřístupňován.

Pro účely testování poskytuje centrální autorita registrátorům možnost využít testovací resolver. **Identifikátory URN:NBN, které lze užít k identifikaci dokumentů, však musejí být vždy přidělovány produkčním resolverem.**

Katalogem se rozumí elektronický, zpravidla online dostupný (OPAC) katalog registrátora, ve kterém mohou být odkazy na jeho digitalizované dokumenty (digitální instance). K odkazování na tyto dokumenty by měl být využit identifikátor URN:NBN. **Metodika využití identifikátorů URN:NBN v katalogu nevyžaduje, pouze doporučuje.**

Registrátor musí své digitalizované dokumenty vlastními silami nebo prostřednictvím služby jiné instituce dlouhodobě uchovávat v digitálním repozitáři a zpřístupňovat v digitální knihovně, nebo mít alespoň záměr vykonávat tyto aktivity v budoucnosti.

Registrátor může být sám provozovatelem digitálního repozitáře a digitální knihovny, pokud tomu tak není, musí mít dohodu s jinými institucemi, jejichž digitální repozitář a digitální knihovnu bude využívat k dlouhodobému uchovávání, resp. zpřístupňování svých digitalizovaných dokumentů. Registrátor musí dodržovat pravidla stanovená touto metodikou. (V případě, že má dohodu s jinými institucemi provozujícími digitální repozitář a digitální knihovnu, musí tyto instituce seznámit s pravidly metodiky a dohodnout s nimi jejich dodržování).

Instituce, která ve svém repozitáři uchovává svěřené digitalizované dokumenty, vystupuje v systému ČIDLO v roli Archivátora.

Digitalizační systém může provozovat sám registrátor, nebo jeho zástupce (digitalizační firma). Digitalizační systém musí být schopen komunikace s resolverem. Produkční resolver přiděluje identifikátory URN:NBN digitalizačnímu systému a získává od něj identifikátorová metadata.

Digitalizační systém vykonává tyto funkce:

- navazuje komunikaci s resolverem za účelem získání identifikátoru URN:NBN,
- dodává do resolveru identifikátorová metadata,
- zajišťuje zápis přiděleného identifikátoru URN:NBN do metadat digitalizovaného dokumentu.

Tabulka 1: Role v systému ČIDLO

Registrátor	Instituce (nebo její část), která je vlastníkem digitálního dokumentu, kterému nechala přidělit identifikátor URN:NBN, nebo která plánuje toto přidělení. Ve většině případů platí, že 1 instituce = 1 registrátor. Registrátor může současně zastávat funkci archivátora nebo provozovatele digitální knihovny.
Centrální autorita	Centrální autoritou v systému ČIDLO je Národní knihovna ČR. Součástí centrální autority jsou zaměstnanci podílející se na jeho provozu, zejména kurátor resolveru a technický správce resolveru, softwarové řešení (resolver) a technické a organizační zajištění.
Kurátor ČIDLO	Role zaměstnance centrální autority, který spravuje obsahovou část resolveru a zajišťuje koordinační knihovnické služby spojené se systémem ČIDLO. Mezi ně patří: rozhodnutí, jaké dokumenty spadají do českého kulturního dědictví, volba kódů pro registrátory, dohody s registrátory, dohody o účasti na systému ČIDLO, poskytování konzultací registrátorům, koordinace postupu registrace registrátorů, aktualizace a urgence údajů od účastníků systémů apod. Email: urnnbn@nkp.cz
Archivátor	Provozovatel digitálního repozitáře, ve které jsou archivovány digitální dokumenty označené identifikátorem URN:NBN v systému ČIDLO. Archivátorem může být sám registrátor, nebo tuto roli může vykonávat jiná instituce na základě dohody s registrátorem.
Provozovatel digitální knihovny	Role instituce v systému, která zajišťuje zpřístupňování dokumentů s identifikátorem URN:NBN.
Katalog	OPAC, katalog registrátora.

8.1 Návod pro zápis URN:NBN identifikátorů do elektronického katalogu

V případě využití identifikátoru URN:NBN v elektronickém katalogu je třeba dodržovat jeden z doporučených zápisů podle pravidel AACR2, případně podle chronologicky mladších pravidel užívaných v ČR.

Identifikátor URN:NBN by měl být zanesen do katalogu k bibliografickému záznamu předlohy digitálního dokumentu. Zápis identifikátoru se řídí katalogizačními pravidly AACR2 a je zapisován dle kódovníku MARC 21.

Pro zápis identifikátoru URN:NBN jsou určena pole 024 a 856. Primární je pole 024.

Pole 024

Nastavení indikátorů:

První indikátor: lze nastavit indikátor 7, v případě nutnosti může být použit i indikátor 8.

Druhý indikátor: typicky bude nastavena hodnota #

V případě nastavení prvního indikátoru 7, musí následovat podpole 2 (odkaz na číselník s dalšími identifikátory včetně URN:NBN). Dále je třeba využít podpole a, do kterého se zapisuje vlastní kód či standardní číslo, nyní tedy URN:NBN.

Př.

024 7# \$a urn:nbn:cz:nk001-0004su \$2 urn

Pole 856

První indikátor: Standardně by měl být vyplněn ukazatel 4, případně je přípustný, byť neodpovídající, #.

Druhý indikátor: Měl by být nastaven na hodnotu 1, případně je přípustný, byť opět neodpovídající, indikátor #.

Podpole: lze užít jen podpole u.

Př.

856 41 \$u <https://resolver.nkp.cz/urn:nbn:cz:nk-0004su>

Doporučení pro zápis do OPAC lze beze zbytku využít pro identifikaci digitalizovaných monografií a jejich svazků. Pokud by měl být využit i v případě seriálů, je nutné, aby konkrétní registrátoři konzultovali s kurátorem systému, jakou úroveň chtějí odkazovat. V takovém případě lze využít i šablonu PERIODICAL (viz kapitola 11).

9. PŘEDMĚT IDENTIFIKACE

Systém ČIDLO je určen pro identifikaci digitalizovaných dokumentů.

Digitalizovaný dokument je množina souborů v určitém formátu, která jako celek, spolu se strukturálními, administrativními a popisnými metadaty, tvoří při zobrazení v příslušné aplikaci (např. internetový prohlížeč) obraz celé tištěné předlohy včetně zachování posloupnosti stran. Tištěnou předlohou se rozumí předloha na úrovni provedení podle FRBR. To znamená, že na obrazové odlišnosti mezi jednotlivými exempláři téže tištěné předlohy (např. exempláři jednoho vydání knihy) se nehledí.

Strukturální metadata popisují vztahy mezi jednotlivými soubory, z nichž se skládá digitalizovaný dokument, a vztah těchto souborů k jednotlivým částem tištěné předlohy. Digitalizovaný dokument vznikl jako konečný výstup skenování tištěné předlohy v rámci jednoho digitalizačního projektu (tj. skenování tištěných předloh a jejich zpracování podle konkrétní specifikace datových a metadatových formátů) a je vlastněn jedním vlastníkem (instituce, která je vlastníkem exempláře tištěné předlohy).

Odlišují se dvě verze digitalizovaného dokumentu - archivní dokument a uživatelský dokument. Archivní i uživatelský dokument mají tentýž obsah (obraz tištěné předlohy), vlastníka a původ (vznikly v rámci téže produkce), mohou se lišit pouze konkrétním obrazovým formátem nebo např. použitým kompresním poměrem v rámci jednoho formátu.

Archivní dokument je tvořen archivními kopiemi (tj. soubory v takovém konkrétním grafickém formátu, který je vhodný pro dlouhodobé uchování v digitálním úložišti) a odpovídajícími strukturálními metadaty. Uživatelský dokument je tvořen uživatelskými kopiemi (tj. soubory v takovém konkrétním grafickém formátu, který je vhodný pro zpřístupňování uživatelům v digitální knihovně).

Archivní dokument ve formátu, ve kterém byl vytvořen při produkci, se nazývá původní archivní dokument. Tento původní archivní dokument je následně dodán do digitálního repozitáře, kde se posléze stává součástí první verze balíčku AIP.

V budoucnosti dojde v digitálním úložišti k tomu, že za účelem dlouhodobého uchování bude původní archivní dokument převeden v rámci formátové migrace do nového formátu a uložen jako součást nové verze balíčku AIP. Takto vzniklý archivní dokument se nazývá nástupnický archivní dokument (první generace). Tento proces se může později opakovat a budou vznikat nové nástupnické archivní dokumenty mladších generací, uložené v dalších nových verzích balíčku AIP. Aktuálním archivním dokumentem se nazývá takový,

který je obsažen v současné nejnovější verzi balíčku AIP (může jím tedy být buď původní archivní dokument, nebo nástupnický archivní dokument jakékoliv generace).

Uživatelský dokument ve formátu, ve kterém byl vytvořen při produkci, se nazývá původní uživatelský dokument. Původní uživatelský dokument je následně dodán do digitální knihovny, kde je uložen a zpřístupňován uživatelům. **Konkrétní URL adresa, ze které může uživatel získat přístup k celému uživatelskému dokumentu, se nazývá digitální instance.** Na této adrese jsou zpravidla uvedena bibliografická metadata popisující obsah uživatelského dokumentu a odkazy na jednotlivé části uživatelského dokumentu (obrázky), jejichž postupným prohlížením může uživatel získat obraz celé tištěné předlohy.

V budoucnosti může dojít k tomu, že z důvodu změny uživatelských požadavků bude digitální úložiště požádáno, aby formátovou migrací vytvořilo novou verzi uživatelského dokumentu z aktuálního archivního dokumentu. Tato nová verze uživatelského dokumentu, odvozená formátovou migrací z aktuálního archivního dokumentu, se nazývá nástupnický uživatelský dokument.

Digitální knihovna může navíc sama, podle svého uvážení, vytvářet další verze uživatelského dokumentu formátovou migrací z původního uživatelského dokumentu nebo z nástupnického uživatelského dokumentu. Všechny takové verze se souhrnně označují jako jiné verze uživatelského dokumentu. Tyto verze nejsou původním uživatelským dokumentem, protože nebyly vytvořeny při produkci, ani nástupnickým uživatelským dokumentem, protože nebyly vytvořeny formátovou migrací v digitálním úložišti.

9.1 Model dokumentu v systému ČIDLO

Model identifikovaného dokumentu je založen na triadickém systému: intelektuální entita – digitální reprezentace/dokument – digitální instance.

9.1.1 Intelektuální entita

Intelektuální entita je jednotka intelektuálního obsahu, která je jako jedna jednotka zpracovávána v informačních procesech (katalogizace, správa, zpřístupňování, archivace apod.). V případě digitalizovaných dokumentů je intelektuální entitou vždy tištěná předloha na úrovni provedení (nebo jeho části).

Závazná pravidla:

- registrátor musí být schopen k této jednotce intelektuálního obsahu dodat bibliografické informace, které popisují právě tuto jednotku (nikoliv jen její část, nikoliv jen celek, jehož je součástí)
 - tyto informace musí povinně obsahovat alespoň název
- intelektuální entita představuje smysluplný informační celek
 - vyloučeny jsou příliš malé (např. stránka knihy⁵) nebo velké části (např. sbírka knih)
 - možností zpracování je kapitola či článek (analytikum) – viz příspěvek ve sborníku
 - takové části musejí být vždy samostatně popsány (bibliografický popis) a samostatně zpřístupňovány
- intelektuální entita představuje neměnný informační celek
 - vyloučeny jsou dynamické celky (jako např. webové stránky – platí pro možné uplatnění v budoucnosti pro digitálně vzniklé dokumenty)
- intelektuální entita v případě digitalizovaných dokumentů je vždy takový obsah (braz tištěné předlohy), který je stejný na úrovni provedení (manifestation) podle modelu FRBR⁶
 - odlišnosti v rámci jednotlivých exemplářů (jednotek podle FRBR) se neberou v potaz
 - výjimka – specifický exemplář (např. opatřený textovým komentářem vlastníka exempláře vepsaným do knihy – filozofická kniha opatřená poznámkami T. G. Masaryka), pokud je specificky popsán bibliografickými metadaty
- intelektuální entita musí představovat informační celek, který má trvalou hodnotu a význam („*lasting value and significance*“ - definice digitálního dědictví podle organizace UNESCO⁷)
 - vyloučeny jsou typy intelektuálních entit, které se podle obecného mínění nepovažují za kulturní dědictví (např. referáty na VŠ, učební pomůcky apod.)
- intelektuální entita má vždy samostatná bibliografická metadata.

⁵ Jiný případ samozřejmě je, pokud jedna stránka reprezentuje např. jednu povídku (viz např. Borgesovy povídky)

⁶ http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf [citováno 2014-09-25]

⁷ <http://unesdoc.unesco.org/images/0013/001331/133171e.pdf#page=80>. Citováno [2017-09-25]

9.1.2 Digitální reprezentace/ dokument

Digitální reprezentace je definována jako jeden nebo více počítačových souborů, které digitálně reprezentují jednu intelektuální entitu.

Signifikantní vlastnosti digitálního dokumentu jsou:

- a) vlastník,
- b) existence samostatných technických metadat,
- c) původ (digitalizační projekt),
- d) dědění bibliografických metadat po intelektuální entitě.

Primární verzí digitálního dokumentu je vždy jeho archivní verze, tedy archivní dokument. Všechny ostatní verze (verze vzniklé formátovou migrací, uživatelské dokumenty) se považují za autentické, pokud vizuálně reprezentují tentýž obraz tištěné předlohy, mají téhož vlastníka a tentýž původ (stejná digitalizace). Typicky bude existovat u jednoho registrátora právě jeden dig. dokument reprezentující IE. Každý digitální dokument má přiřazeno právě jedno URN:NBN. Je registrován právě jedním Registrátorem a archivován alespoň jedním Archivátorem. Digitální dokument je tím, co je identifikováno systémem ČIDLO, nikoli Intelektuální entita. Identifikátor musí být přidělen v závěrečné fázi vzniku digitálního dokumentu.

9.1.3 Digitální instance

Digitální instance je lokalizace digitálního dokumentu prostřednictvím tzv. vstupní stránky.

Signifikantní vlastnosti digitální instance jsou:

1. vstupní stránka, která se vyznačuje:
 - a) jedním URL, na němž jsou:
 - b) přístupná bibliografická, volitelně i technická metadata k uživatelskému dokumentu,
 - c) přístupné odkazy na jeden nebo více verzí uživatelského dokumentu,
2. v rámci jedné digitální knihovny existuje pouze jedna vstupní stránka.

K jedné digitální reprezentaci může existovat variabilní počet digitálních instancí. Pokud dojde k digitalizaci téže intelektuální entity dvěma či více institucemi, pak na základě rozpoznání vlastníka jako signifikantní hodnoty v popisu dokumentu vznikne více záznamů digitálních dokumentů. U každého dokumentu budou uvedeny všechny možné digitální instance, tzn. všechny lokace, na kterých bude daný dokument s daným identifikátorem dostupný.

Příklad: Pokud NK ČR zdigitalizuje určitý titul a na základě právní dohody ho poskytne Moravské zemské knihovně, aby ho zpřístupnila ve své digitální knihovně (bude tak i nadále označen pomocí URN:NBN identifikátoru s kódem nk), tak v resolveru budou dostupné dvě digitální instance připojené k jednomu záznamu dokumentu. Pokud však MZK a NK ČR zdigitalizují tentýž titul (tedy duplicitně), budou přiděleny dva URN:NBN identifikátory, záznam dokumentu bude v resolveru vložen dvakrát a ke každému bude připojena jedna digitální instance.

9.1.4 Vzájemné vztahy

Jedna intelektuální entita může být reprezentována jak pouze jedním digitálním dokumentem, tak více digitálními dokumenty. Jeden digitální dokument musí být lokalizován prostřednictvím jedné nebo více digitálních instancí.



Obrázek 3: Popis vztahů mezi intelektuální entitou a digitální reprezentací.

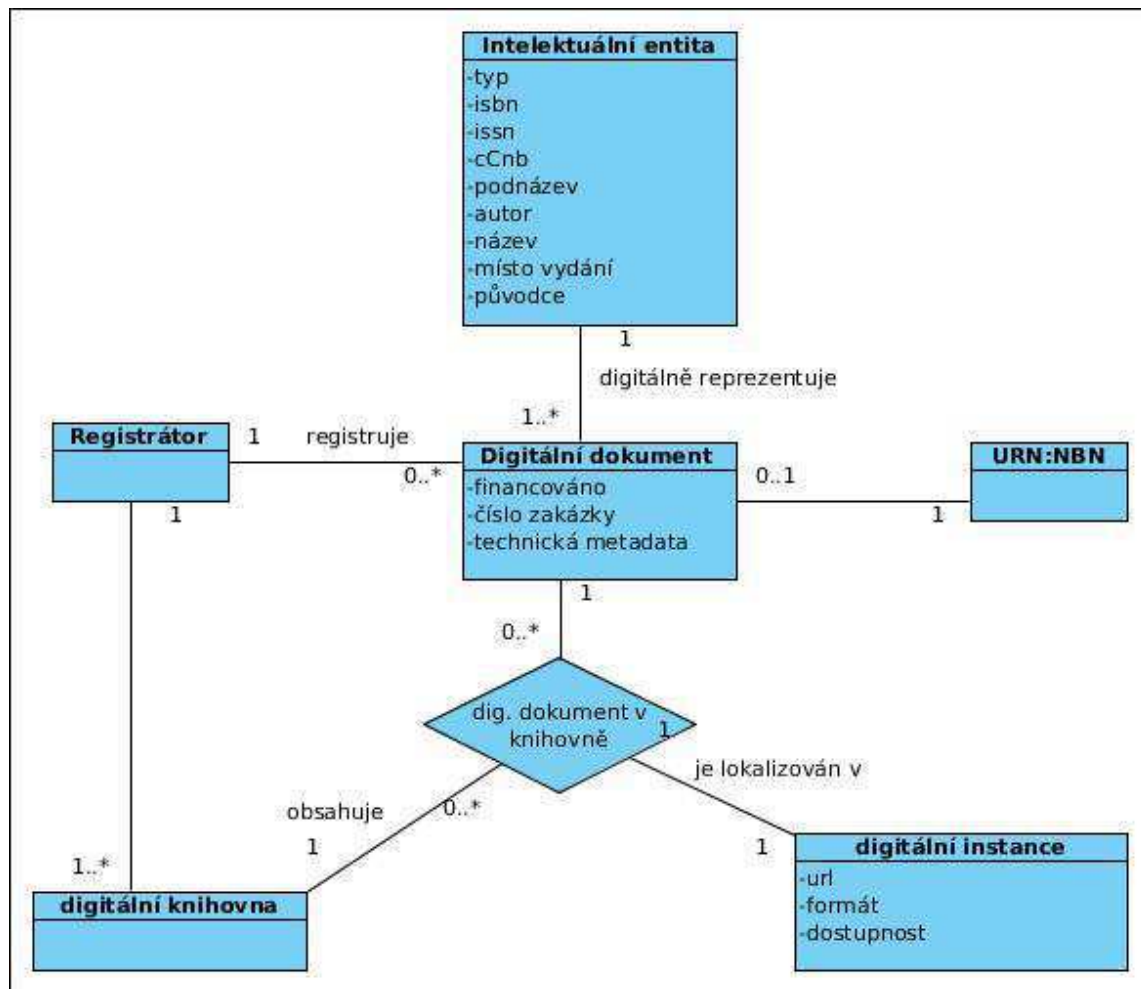
9.1.5 Datový model

Z výše popsaného modelu dokumentu vychází i datový model systému ČIDLO, kdy jádrem celého systému je digitální reprezentace v rámci terminologie systému ČIDLO označovaná jako digitální dokument (digitalDocument). Digitální dokument je registrátorem registrován a je mu přidělen URN:NBN identifikátor. Ten označuje konkrétní množinu počítačových souborů. Identifikátor URN:NBN slouží k identifikaci digitálních dokumentů, nikoliv jejich analogových předloh. Pokud z jednoho titulu budou vytvořeny dva zdánlivě identické digitální dokumenty (stejná obrazová data), ale jejich producenty budou jiní vlastníci, jsou v tomto pojetí chápány jako dva odlišné dokumenty (mezi signifikantní vlastnosti patří vlastníci, který je odvozen od vlastníka předlohy).

Bibliografický popis tedy zůstává na úrovni intelektuální entity, zatímco technická

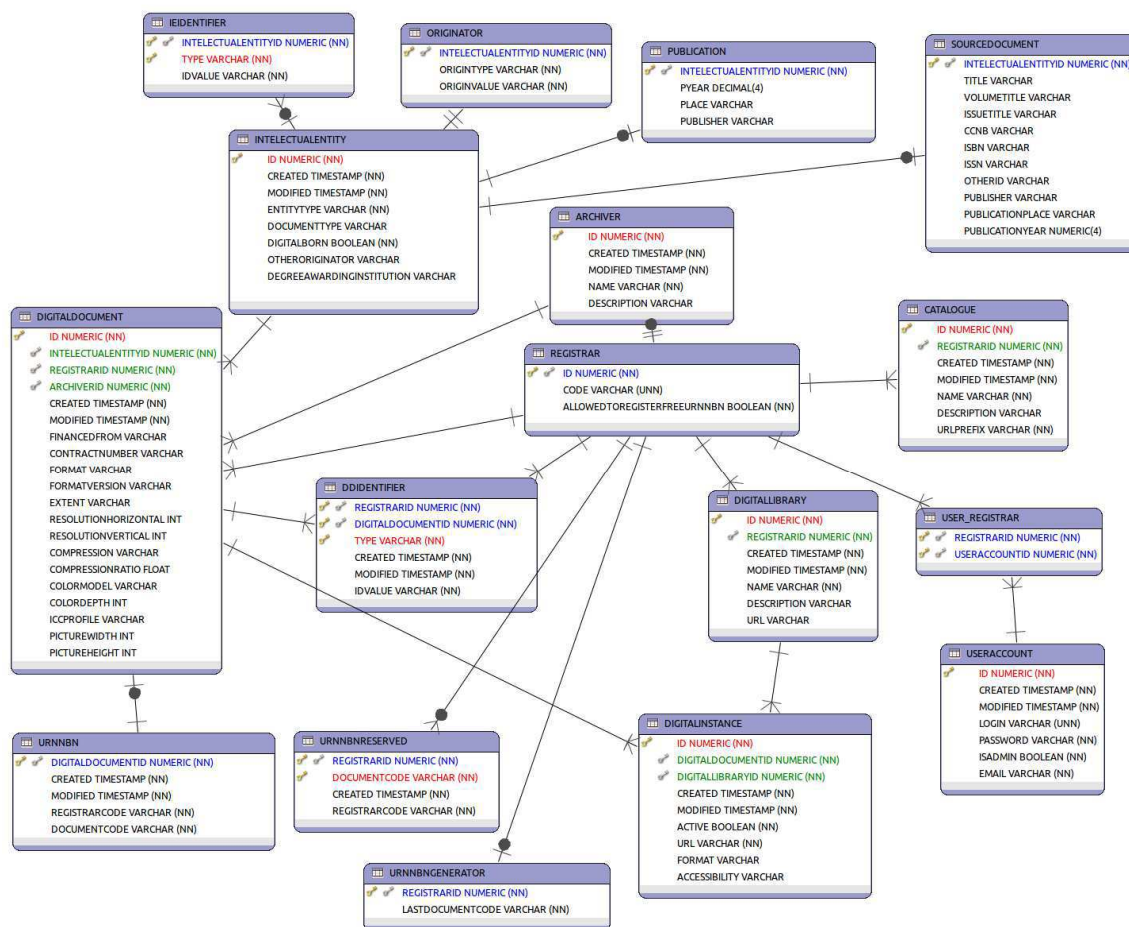
metadata, vlastní identifikátor a administrativní popis se vážou k digitální reprezentaci (dokumentu).

Koncepční model vztahů mezi základními objekty je zobrazen na obrázku č. 4.



Obrázek 4: Koncepční model objektů v systému ČIDLO.

Všechny databázové tabulky a vztahy mezi nimi zobrazuje fyzický model na obrázku č. 5.



Obrázek 5: Vztahy mezi tabulkami v databázi.

9.2 Vznik a struktura digitalizovaného dokumentu

Digitalizovaný dokument musí mít podobu, definovanou ve standardech digitalizace NDK. Zejména to znamená, že musí používat:

- zde definovanou metadatovou specifikaci (vybrané položky konkrétních metadatových formátů),
- doporučení pro datové formáty (specifikace formátu JP2 pro archivní a uživatelské kopie),
- strukturální metadata
- obsah digitalizovaných dokumentů musí odpovídat stanoveným základním intelektuálním entitám (monografie, periodikum apod.), resp. jejich povolené granularitě (vyšší / nižší intelektuální entita; např. článek periodika, kapitola monografie).

V případě nestandardního dokumentu platí, že musí mít metadatovou strukturu umožňující korektní propojení s resolverem a zároveň by měl být uložen v důvěryhodném dlouhodobém úložišti. Identifikátor URN:NBN lze přidělit digitalizovanému dokumentu, který obsahově odpovídá jakékoliv předepsané intelektuální entitě, vyjma titulu periodika. Různým úrovním granularity se přidělují jiné identifikátory URN:NBN, takže například digitalizovaný dokument, jehož obsah odpovídá článku periodika X, dostane jeden identifikátor URN:NBN, a digitalizovaný dokument, který odpovídá číslu periodika X, dostane jiný identifikátor URN:NBN. Skutečnost, že identifikátor URN:NBN v prvním případě identifikuje část dokumentu identifikovaného druhým identifikátorem, nelze ze samotného identifikátoru URN:NBN nijak zjistit.

Digitalizovaný dokument je součástí balíčku PSP, který obsahuje mimo jiné obě obrazové verze digitalizovaného dokumentu současně (tj. uživatelský a archivní dokument). Balíček PSP je tvořen hlavním adresářem s několika podadresáři a soubory v grafických a textových formátech (podle platných Standardů digitalizace NDK to jsou JPEG2000, TXT, MD5 a XML), které se nalézají v hlavním adresáři a jeho podadresářích. Obsah PSP balíčku odpovídá definovanému konceptu původního archivního a uživatelského dokumentu.

10. POSTUP PŘIDĚLOVÁNÍ IDENTIFIKÁTORŮ, PRAVIDLA SYSTÉMU

ČIDLO

Systém trvalé identifikace založené na URN:NBN identifikátorech předpokládá spolupráci různých institucí. Existují různé role v systému: centrální autorita, registrátor, archivátor a digitální knihovna. **Agregátoři, tedy instituce, které data nevlastní, ale pouze je zpřístupňují, nebo je u sebe ukládají duplicitně, nemohou žádat o přidělení URN:NBN identifikátoru.**

Registrátor musí pro zapojení do systému ČIDLO provést následující kroky.

10.1 Obecné povinnosti

Registrátor je povinen po celou dobu své aktivity dodržovat pravidla stanovená touto metodikou a dodržovat předpisy popisující identifikovaný dokument, jak byla popsána v kapitole 9. Jednou zasláný záznam registrovaného dokumentu je možné měnit jen za vymezených podmínek popsaných v kapitole 14. V obecné rovině musí registrátor splnit zejména tyto požadavky:

- a) identifikovaný digitální dokument je nutné trvale archivovat (v rámci digitálního repozitáře provozovaného registrátorem nebo jinou institucí, se kterou má registrátor dohodu o archivaci),
- b) identifikovaný dokument musí být trvale zpřístupňován v internetové síti (v rámci své vlastní digitální knihovny, nebo digitální knihovny provozované jinou institucí, s níž má registrátor dohodu o zpřístupňování svých dokumentů). V rámci digitální knihovny musí být identifikovaný dokument vystaven tak, aby **byl identifikátor URN:NBN viditelný** a uvedený v sekci bibliografických údajů popisujících daný dokument,
- c) Registrátor musí v digitální knihovně identifikátor URN:NBN namapovat do metadat pro svůj OAI-provider,
- d) **žádost o přidělení identifikátoru musí digitalizační systém zaslat ve fázi dokončování produkčního balíčku, o identifikátor nelze žádat později.**

Z daných pravidel existují výjimky možné po dohodě s kurátorem systému:

- a) registrované dokumenty nemusí být zveřejněné ihned, ale v určité časové lhůtě,
- b) i po vzniku dokumentů je možné mu zpětně přidělit identifikátor, ale toto je možné jen

za situace, kdy dokumenty byly již vytvořeny, ale dosud nezveřejněny

10.2 Povinnosti před přidělením URN:NBN identifikátoru (registrace registrátora)

Registrátorem se může stát každý, kdo je k tomu oprávněn podle pravidel v kapitole 7. Každému žadateli o registraci, který splňuje pravidla, musí být vyhověno. Ve fázi vymezené okamžikem žádosti o registraci registrátora, kterou žadatel provede zpravidla pomocí elektronické pošty na adresu urnnbn@nkp.cz (případně písemně na adresu určeného pracovníka Národní knihovny České republiky) a začátkem produkčního přidělování identifikátorů, je nutné provést následující kroky:

- a) Registrátor oznámí kurátorovi resolveru zájem o přidělování URN:NBN a seznámí se s touto metodikou a jejími pravidly. Dále bude postupovat v souladu s jejími postupy.
- b) Registrátor oznámí kurátorovi:
 - a. Katalog, ve kterém bude uvádět identifikátory URN:NBN (je na volbě registrátora, zda tento údaj bude ve svém OPAC zveřejňovat).
 - b. Název a adresu digitální knihovny, ve které budou digitální dokumenty zpřístupňovány (URL adresa pro OAI-repository) – tato knihovna nemusí být ve správě registrátora, postačí, pokud bude mít s vlastníkem knihovny dohodu. Provozovatel této digitální knihovny musí respektovat pravidla metodiky pro digitální knihovny zapojené do systému ČIDLO.
 - c. Název archivátora (instituce, která bude tyto dokumenty trvale archivovat) – *platí pouze v případě, že archivátorem je jiná instituce, než je registrátor*
 - d. Seznam všech možných sigel v katalogu označujících dokumenty registrátora (pokud je jich víc než jedna), pokud bude takové dokumenty digitalizovat.
 - e. Kontakt na zpracovatele digitalizace, pokud jím není on sám. Takový zpracovatel se musí řídit pravidly této metodiky. Za jejich dodržení je odpovědný registrátor.
- c) Registrátor či jím pověřený zpracovatel si nastaví komunikační rozhraní pro přidělení identifikátoru resolverem (REST). Registrátor je povinen využívat vždy nejnovější dostupnou verzi API, pokud se nedohodne s kurátorem služby jinak. Dále si registrátor nastuduje postupy pro jednotlivé operace tak, jak jsou popsány

na <https://code.google.com/p/czidlo/> Především nastaví správné rozpoznávání druhů digitálních dokumentů systému ČIDLO. V případě malých digitalizací lze registrátorovi udělit práva k ručnímu zadávání požadovaných metadat pomocí webového formuláře.

- d) Registrátor otestuje nastavení svého systému na testovací instanci resolveru, kam dostane přístupová práva od kurátora systému.
- e) Po splnění předcházejících podmínek vytvoří kurátor systému ČIDLO záznam registrátora a uživatelský účet registrátora a předá ho pověřené osobě zastupující registrátora. Dále mu sdělí další potřebné údaje (např. kód archivátora).

10.3 Uživatelské účty

K jednomu uživatelskému účtu mohou být navázána práva spravovat více jmenných prostorů. Zároveň každý jmenný prostor může být spravován z více uživatelských účtů současně. Uživatelský účet může získat i instituce či fyzická osoba, která není registrátorem. Musí však k tomu být ze strany alespoň jednoho registrátora zmocněna. Kurátor systému pak na základě informace od registrátora takový účet neprodleně zřídí. Povolení k produkci dostane pověřená instituce až po otestování svého nastavení pomocí testovací instance systému ČIDLO. Za dodržování pravidel systému uživatelem takového uživatelského účtu je odpovědný registrátor, který ho zmocnil.

Příklad: Dodavatelská firma získala pověření od více registrátorů, aby zpracovávala jejich digitalizaci, dostane přidělen vlastní uživatelský účet, ke kterému jsou přidělena práva k danému jmennému prostoru. Vedle této externí digitalizace však provádí registrátor i vlastní digitalizaci. K té bude používat vlastní účet. V případě online komunikace s resolverem zajišťuje kontrolu duplicitního přidělení identifikátoru sám resolver v okamžiku zpracování registrace dokumentu. V případě využití rezervace resolver kontroluje, zda nebyl identifikátor použit dvakrát a případně zamítne dokončení registrace.

10.3.1 Možnosti operací při správě jmenného prostoru

Uživatel (registrátor či jeho pověřený zástupce) může v rámci svého jmenného prostoru provádět různé operace, které mu slouží pro potřeby úspěšného vykonávání správy nad přiděleným prostorem. Kromě operací, které tvoří jádro funkcí systému ČIDLO a které budou popsány níže, jde především o možnost zobrazit záznam registrátora, který mj. ukáže povolené způsoby registrace dokumentů (viz např. záznam MZK: <http://resolver.nkp.cz/api/v3/registrars/boa001>). Dále lze pomocí jak webového rozhraní, tak

API získat informaci o počtu digitálních dokumentů daného registrátora – ve webovém rozhraní v záložce Procesy nebo prostřednictvím dotazu ve tvaru „[http://resolver.nkp.cz/api/v3/registrars/\(kód registrátora\)/digitalDocuments](http://resolver.nkp.cz/api/v3/registrars/(kód%20registrátora)/digitalDocuments)“. API nabízí i možnost zobrazení informací o stavu a dalších náležitostech konkrétního identifikátoru – tvar requestu je (pro příklad dokumentu s urn:nbn:cz:boa001-000056) „<http://resolver.nkp.cz/api/v3/urnnbn/urn:nbn:cz:boa001-000056>“. Popis všech možných operací, které API resolveru nabízí, je dostupný na projektové stránce <https://code.google.com/p/czidlo/wiki/ApiV3>. O operacích, které může vlastník účtu samostatně provádět, je třeba se vždy dohodnout s kurátorem služby, obvykle však lze mít za to, že ve svém jmenném prostoru může registrátor používat všechny dostupné funkcionality.

10.4 Postup ve fázi přidělení URN:NBN identifikátoru

Po zahájení produkce digitálních dokumentů může registrátor zvolit dva způsoby získávání identifikátorů. Metodikou doporučená je možnost online získávání identifikátorů, tedy ta, kdy registrátor odešle požadované údaje do resolveru a nazpět získá identifikátor. Druhým způsobem je získání identifikátoru cestou rezervace. Registrátor si rezervuje balík identifikátorů (doporučeno je maximálně 1 000 identifikátorů), které následně použije a potvrdí jejich registrací zasláním požadovaných metadat. Je doporučeno, aby registrátoři získali určité množství identifikátorů do zásoby tak, aby jimi mohli pokrýt případný krátkodobý výpadek služby ČIDLO. Na straně registrátora je zvýšené riziko omylu, protože hrozí použití jednou rezervovaného identifikátoru dvakrát. V takovém případě resolver vyhodnotí duplicitu a odmítne dokončit rezervaci.

Registrátor má v této fázi následující povinnosti:

- a) Dodání vybraných bibliografických a technických metadat dokumentu resolveru výměnou za přidělení identifikátoru,
- b) přidělení téhož URN:NBN současně do metadat všech verzí téhož dokumentu (obvykle jde o dvě verze – archivní a uživatelský dokument [AIP a DIP]).

Doporučený zápis do standardů MODS / Dublin Core:

MODS

```
<mods:identifier type="urnnbn">urn:nbn:cz:nk-000h4u</mods:identifier>
```

DC

Národní knihovna České republiky, Klementinum 190,110 00 Praha 1
www.nkp.cz; www.ndk.cz; <https://resolver.nkp.cz>

<dc:identifier>urnnbn:urn:nbn:cz:nk-000h4u</dc:identifier>

Zápis musí být realizován na závěr produkce PSP balíčku.

Registrátor je nadále oprávněn spravovat svůj jmenný prostor, tedy může v rámci pravidel řízení životního cyklu dokumentu (kapitola 14 metodiky) provádět úpravy dokumentů a digitálních instancí. Především může provádět deaktivace dokumentů, jejich slučování a rozpojování, aktualizovat údaje o digitálních instancích přes webové rozhraní nebo pomocí OAI sklizně.

10.4.1 Příklady operací při registraci digitálního dokumentu

URN:NBN identifikátor je přidělován digitálnímu dokumentu operací *Registrace digitálního dokumentu*. Vstupem operace je metadatový záznam dokumentu a také záznam IE, kterou dig. dokument reprezentuje. Vstupní data jsou pak validována a pokud v nich není nalezena chyba, je v systému vytvořena nová IE a její digitální dokument. Výstupem operace je pak přiřazený URN:NBN identifikátor. Pokud resolver nalezne chybu, zachová se podle jejího charakteru:

1) jde-li o nepřítomnost povinného údaje, odmítne provést registraci,

2) jde-li o chybu v hodnotě některého z kontrolovaných polí (typicky číslo České národní bibliografie nebo ISBN), bude registrace provedena, ale chybná hodnota nebude uložena.

Registrovaný dokument musí být validní podle:

<http://resolver.nkp.cz/api/v3/digDocRegistration.xsd>

Schéma pro import dokumentu je dostupné zde:

<http://resolver.nkp.cz/api/v3/digInstImport.xsd>

V následujících příkladech jsou využity ukázky konkrétních dokumentů. Jednotlivá pole v databázi budou vysvětlena v kapitole 11 a to včetně doporučení jakými hodnotami tato pole lze plnit.

1. Registrace digitálního dokumentu - mód BY_RESOLVER

CZIDLO přiřadí volné (nerezervované) URN:NBN.

Request

```
POST http://resolver.nkp.cz/api/v3/registrars/boa001/digitalDocuments
```

V těle http dotazu se nachází tento obsah (Tělo requestu bude obsahovat importní xml s metadaty intelektuální entity):

```

<?xml version="1.0" encoding="UTF-8"?>
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  <r:monograph>
    <r:titleInfo>
      <r:title>Babička</r:title>
      <r:subTitle>Obrazy z venkovského života</r:subTitle>
    </r:titleInfo>
    <r:ccnb>cnb002251177</r:ccnb>
    <r:isbn>8090119964</r:isbn>
    <r:otherId>DOI:TODO</r:otherId>
    <r:documentType>kniha</r:documentType>
    <r:digitalBorn>>false</r:digitalBorn>
    <r:primaryOriginator type="AUTHOR">Božena
Němcová</r:primaryOriginator>
    <r:otherOriginator>Adolf Kašpar</r:otherOriginator>
    <r:publication>
      <r:publisher>Československý spisovatel</r:publisher>
      <r:place>V Praze</r:place>
      <r:year>2011</r:year>
    </r:publication>
  </r:monograph>
  <r:digitalDocument>
    <r:archiverId>3</r:archiverId>
    <r:registrarScopeIdentifiers>
      <r:id type="OAI_harvester">1234</r:id>
      <r:id type="signatura">PK-0123.999</r:id>
      <r:id type="K4_pid">uuid:38bfc69d-1d10-4822-b7a4-
c7531bb4aedg</r:id>
    </r:registrarScopeIdentifiers>
    <r:financed>norské fondy</r:financed>
    <r:contractNumber>123</r:contractNumber>
    <r:technicalMetadata>
      <r:format version="1.0">jpeg</r:format>
      <r:extent>245 x jpeg2000;245 x mods-alto;1 x mods</r:extent>
      <r:resolution>
        <r:horizontal>1280</r:horizontal>
        <r:vertical>1024</r:vertical>
      </r:resolution>
      <r:compression ratio="0.3">LZW</r:compression>
      <r:color>
        <r:model>RGB</r:model>
        <r:depth>24</r:depth>
      </r:color>
      <r:iccProfile>some ICC profile</r:iccProfile>
      <r:pictureSize>
        <r:width>600</r:width>
        <r:height>1000</r:height>
      </r:pictureSize>
    </r:technicalMetadata>
  </r:digitalDocument>
</r:import>

```

Response

```

<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <urnNbn>
    <status>ACTIVE</status>
    <value>urn:nbn:cz:boa001-000094</value>

```

```
<countryCode>cz</countryCode>
<registrarCode>boa001</registrarCode>
<documentCode>000094</documentCode>
<digitalDocumentId>200</digitalDocumentId>
<registered>2013-01-18T00:27:06.424+01:00</registered>
</urnNbn>
</response>
```

2. Registrace digitálního dokumentu - mód BY_RESERVATION

Nejdříve je nutné provést rezervaci balíku URN:NBN identifikátorů, které budou následně registrovány. Žádost o rezervaci vypadá takto:

Request

```
POST
http://resolver.nkp.cz/api/v3/registrars/boa001/urnNbnReservations?size=5
```

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <urnNbnReservation>
    <urnNbn>urn:nbn:cz:boa001-000006</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000007</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000008</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000009</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-00000a</urnNbn>
  </urnNbnReservation>
</response>
```

Seznam rezervovaných URN:NBN získáte tímto dotazem:

Request

```
GET http://resolver.nkp.cz/api/v3/registrars/boa001/urnNbnReservations
```

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <urnNbnReservations>
    <maxReservationSize>100</maxReservationSize>
    <defaultReservationSize>10</defaultReservationSize>
    <reserved totalSize="8">
      <urnNbn reserved="2012-10-
03T00:22:06.930+02:00">urn:nbn:cz:boa001-00000c</urnNbn>
      <urnNbn reserved="2012-10-
03T00:22:07.051+02:00">urn:nbn:cz:boa001-00000d</urnNbn>
      <urnNbn reserved="2012-10-
03T00:22:07.187+02:00">urn:nbn:cz:boa001-00000e</urnNbn>
      <urnNbn reserved="2012-10-
03T01:09:41.566+02:00">urn:nbn:cz:boa001-00000f</urnNbn>
```

```
        <urnNbn reserved="2012-10-
03T01:09:41.632+02:00">urn:nbn:cz:boa001-00000g</urnNbn>
        <urnNbn reserved="2012-10-
03T01:09:41.699+02:00">urn:nbn:cz:boa001-00000h</urnNbn>
        <urnNbn reserved="2012-10-
03T01:09:41.765+02:00">urn:nbn:cz:boa001-00000i</urnNbn>
        <urnNbn reserved="2012-10-
03T01:09:41.832+02:00">urn:nbn:cz:boa001-00000j</urnNbn>
    </reserved>
</urnNbnReservations>
</response>
```

Rezervované URN:NBN je následně nutné spojit s digitálním dokumentem a aktivovat:

Request

POST <http://resolver.nkp.cz/api/v3/registrars/boa001/digitalDocuments>

V těle http dotazu se nachází tento obsah:

```
<?xml version="1.0" encoding="UTF-8"?>
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  <r:monograph>
    <r:titleInfo>
      <r:title>Babička</r:title>
      <r:subTitle>Obrazy z venkovského života</r:subTitle>
    </r:titleInfo>
    <r:ccnb>cnb002251177</r:ccnb>
    <r:isbn>8090119964</r:isbn>
    <r:otherId>DOI:TODO</r:otherId>
    <r:documentType>kniha</r:documentType>
    <r:digitalBorn>>false</r:digitalBorn>
    <r:primaryOriginator type="AUTHOR">Božena
Němcová</r:primaryOriginator>
    <r:otherOriginator>Adolf Kašpar</r:otherOriginator>
    <r:publication>
      <r:publisher>Československý spisovatel</r:publisher>
      <r:place>V Praze</r:place>
      <r:year>2011</r:year>
    </r:publication>
  </r:monograph>
  <r:digitalDocument>
    <r:archiverId>3</r:archiverId>
    <r:urnNbn>
      <r:value>urn:nbn:cz:boa001-00003b</r:value>
      <r:predecessor value="urn:nbn:cz:boa001-000030" note="spojeno v
jeden dokument"/>
      <r:predecessor value="urn:nbn:cz:boa001-00003a" note="spojeno v
jeden dokument"/>
    </r:urnNbn>
    <r:registrarScopeIdentifiers>
      <r:id type="OAI_harvester">3234</r:id>
      <r:id type="signatura">PK-0123.666</r:id>
      <r:id type="K4_pid">uuid:38bfc69d-1e10-4822-b7a4-
c7531bb4aedg</r:id>
    </r:registrarScopeIdentifiers>
    <r:financed>norské fondy</r:financed>
    <r:contractNumber>123</r:contractNumber>
    <r:technicalMetadata>
```

```

<r:format version="1.0">jpeg</r:format>
<r:extent>245 x jpeg2000;245 x mods-alto;1 x mods</r:extent>
<r:resolution>
  <r:horizontal>1280</r:horizontal>
  <r:vertical>1024</r:vertical>
</r:resolution>
<r:compression ratio="0.3">LZW</r:compression>
<r:color>
  <r:model>RGB</r:model>
  <r:depth>24</r:depth>
</r:color>
<r:iccProfile>some ICC profile</r:iccProfile>
<r:pictureSize>
  <r:width>600</r:width>
  <r:height>1000</r:height>
</r:pictureSize>
</r:technicalMetadata>
</r:digitalDocument>
</r:import>

```

Response

```

<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <urnNbn>
    <status>ACTIVE</status>
    <value>urn:nbn:cz:boa001-00003b</value>
    <countryCode>cz</countryCode>
    <registrarCode>boa001</registrarCode>
    <documentCode>00003b</documentCode>
    <digitalDocumentId>105</digitalDocumentId>
    <reserved>2012-11-20T02:12:07.123+01:00</reserved>
    <registered>2012-11-25T02:09:02.297+01:00</registered>
    <predecessor value="urn:nbn:cz:boa001-000030"/>
    <predecessor value="urn:nbn:cz:boa001-00003a"/>
  </urnNbn>
</response>

```

3. Múd BY_REGISTRAR

Vedle obou zmíněných postupů existuje i možnost přidělování registrátorem, kdy on volí podobu identifikátoru, ale tato není běžně pro registrátory dostupná. Při běžné produkci by s touto možností neměl registrátor přijít do styku.

Příčemž samotná rezervace je provedena následovně:

Rezervace balíku URN:NBN

Request

```

POST
http://resolver.nkp.cz/api/v3/registrars/boa001/urnNbnReservations?size=5

```

Response


```

<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <urnNbnReservation>
    <urnNbn>urn:nbn:cz:boa001-000006</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000007</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000008</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-000009</urnNbn>
    <urnNbn>urn:nbn:cz:boa001-00000a</urnNbn>
  </urnNbnReservation>
</response>

```

10.5 Povinnosti z hlediska zpřístupňování dokumentu

Registrátor je povinen registrované dokumenty vystavit ve své nebo ve spolupracující digitální knihovně. Metodika předpokládá, že se tak stane neprodleně, ale pokud tomu technické či organizační důvody brání, lze tuto lhůtu po konzultaci s kurátorem systému posunout. Registrátor vkládá do záznamu digitálního dokumentu odkaz na digitální instanci v okamžiku zveřejnění v digitální knihovně pomocí automatizované funkce nebo sám či s pomocí kurátora provede v dohodnutém časovém intervalu (alespoň jednou za tři měsíce pokud probíhá produkce, nebo alespoň jednou za šest měsíců, pokud nedochází k importům dokumentů; ví-li registrátor, že došlo ke změně lokace digitálních instancí jeho knihovny, iniciuje sklizeň neprodleně) OAI sklizeň své digitální knihovny. Při zveřejňování identifikovaných dokumentů je třeba dodržet tato pravidla:

- a) trvalé zpřístupňování digitálního dokumentu v internetové síti je v odpovědnosti instituce provozující danou digitální knihovnu, ve které je digitální dokument zpřístupňován čtenáři,
- b) toto zpřístupňování se děje prostřednictvím tzv. vstupní stránky, což je stránka, která má URL adresu a na níž jsou dostupná metadata vztahující se k dokumentu a odkazy na všechny jednotlivé části dokumentu (pokud se například digitální dokument skládá ze 100 souborů ve formátu JPEG, pak musí být možné se ze vstupní stránky na všechny tyto soubory dostat),
- c) digitální knihovna musí provozovat OAI-repository, aby mohl resolver přes tuto službu sklízet aktuální URL adresy dokumentů.

10.5.1 Příklad operace vložení digitální instance

Pro vložení digitální instance se volá POST metoda http requestu v následujícím tvaru:

```
<host_name>:<host_port>/api/resolver/{URN_NBN}/digitalInstances
```

Kde

{URN_NBN} je přidělené URN:NBN

Příklad volání:

<http://resolver.nkp.cz/api/resolver/urn:nbn:cz:tst01-000001/digitalInstances>

Tělo requestu bude obsahovat importní xml s informacemi o digitální instanci. Záznamy vhodné k importu musí projít validací podle schématu digitalInstanceImport.xsd.xml

Elementy importního xml:

1. digitalInstance – instance, která se má naimportovat
2. url – vlastní url, kde je digitální instance přístupná
3. digitalLibraryId – ID digitální knihovny, vyplněno dle digitální knihovny (Krameria; ale lze spolupracovat s jakoukoli digitální knihovnou splňující požadavky), do kterého bude balíček uložen
4. format – formát, v kterém je digitální instance přístupná. Hodnoty jp2; pdf
5. accessibility – přístupnost. Bude nastaveno dle hodnoty nastavené v průběhu digitalizačního WF

Request

POST <http://resolver.nkp.cz/api/v3/resolver/urn:nbn:cz:boa001-000001/digitalInstances>

Tělo dotazu pak obsahuje:

```
<?xml version="1.0" encoding="UTF-8"?>
<digitalInstance xmlns="http://resolver.nkp.cz/v3/">
  <url>http://kramerius3.mzk.cz/kramerius/handle/BOA001/935239</url>
  <digitalLibraryId>4</digitalLibraryId>
  <format>jpg;pdf</format>
  <accessibility>volně přístupné</accessibility>
</digitalInstance>
```

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
  http://resolver.nkp.cz/api/v3/response.xsd">
  <digitalInstance id="52" active="true">
    <url>http://kramerius3.mzk.cz/kramerius/handle/BOA001/935239</url>
    <format>jpg;pdf</format>
    <accessibility>volně přístupné</accessibility>
    <digitalLibraryId>4</digitalLibraryId>
    <created>2013-01-17T02:09:50.349+01:00</created>
  </digitalInstance>
</response>
```

Digitální instance lze vkládat a deaktivovat jak automatizovaně pomocí API, tak manuálně pomocí webového rozhraní.

Pomocí API lze získat informaci o všech digitálních instancích daného dokumentu - <http://resolver.nkp.cz/api/v3/resolver/urn:nbn:cz:boa001-000001/digitalInstances>

10.6 Chybové stavy

Při komunikaci s resolver může docházet k zaslání chybných údajů. Na jednotlivé typy chyb reaguje resolver následujícími kódy.

Tabulka 2: Přehled chybových kódů v „Response“

HTTP kód	Kód chyby v "Response"	Popis
405	METHOD_NOT_PERMITTED	Nepovolená metoda pro danou URL
400	MISSING_PARAMETER	Chybějící povinný parametr, nebo je parametr nutný vzhledem k hodnotám jiných parametrů
400	INVALID_REGISTRAR_CODE	Kód registrátora není validní
400	INVALID_DATA	Tělo http dotazu není well-formedxml, nebo není validní podle import.xsd
400	INVALID_URN_NBN	Urn:nbn má špatnou syntaxi, nebo obsahuje špatný kód registrátora, nebo registrátor nemá povoleno vkládat předem nerezervovaná urn:nbn, přesto bylo urn:nbn ve stavu FREE nalezeno v importu
400	INVALID_ARCHIVER_ID	Id archivátor nalezené v importním xml neoznačuje žádného existujícího archivátora
401	NOT_AUTHENTICATED	Operace si vyžaduje autentizaci, uživatel však není přihlášen
401	NOT_AUTHORIZED	Uživatel je přihlášen, nemá dostatečná práva – není správcem registrátora, který registruje daný objekt
403	FORBIDDEN	Metoda není pro danou

		instanci aplikace povolena. Může se jednat o dočasné řešení (zakázání importů), nebo se může jednat o read-only instanci aplikace určenou pouze pro resolvování a zobrazování záznamů
404	UNKNOWN_REGISTRAR	Registrátor s tímto kódem neexistuje.

11. TYPY DOKUMENTŮ

Typ dokumentu je určen **povinným atributem *intellectualEntityType***, který obsahuje některou z hodnot **MONOGRAPH**, **MONOGRAPH_VOLUME**, **PERIODICAL**, **PERIODICAL_VOLUME**, **PERIODICAL_ISSUE**, **THESIS**, **ANALYTICAL** nebo **OTHER**. Přičemž typ PERIODICAL pro titul seriálu je administrativně zakázáno používat, lze změnit na základě odůvodněné žádosti registrátora.

K upřesnění typu dokumentu v lidsky čitelné podobě slouží atribut *documentType*. Jeho využití je vhodné, například pokud chceme u typu MONOGRAPH rozlišit, zda se jedná o knihu, rukopis nebo starý tisk. **Nutné je upřesnění druhu používat u typu OTHER, zde se můžou objevit hodnoty jako “mapa”, “grafika”.**

Navíc některé atributy samotné tabulky *IntellectualEntity* jsou použitelné jen pro vybrané typy IE. Konkrétně se jedná o atributy *digitalBorn* a *degreeAwardingInstitution*. *DigitalBorn* je použit pro všechny publikovatelné typy IE. Je to binární příznak, který označuje, jestli dílo původně vzniklo v digitální podobě. *DegreeAwardingInstitution* používá pouze typ vysokoškolská práce (THESIS). Atribut slouží k identifikaci vysoké školy.

Metadatový popis může obsahovat několik URN:NBN identifikátorů. Děje se tak v případě, že registrátor chce využít více typů dokumentů, typicky pro číslo seriálu a ročník seriálu. V resolveru vzniknou dva nezávislé záznamy, přičemž za primární se považuje záznam čísla. Tento případ může nastat ještě u dokumentu typu ANALYTICAL. URN:NBN identifikátor získaný pro tuto intelektuální entitu musí být zapsán do metadatového popisu na úroveň předpokládanou Standardy digitalizace NDK pro tento typ dokumentu. Naopak svazek monografie je chápán jako samostatná intelektuální entita a každý je registrován samostatně. Záznam celého titulu knižní řady se nevytváří. K určení **typu dokumentu (*intellectualEntityType*)** lze využít výsledek **transformace záznamu v MARC21 do standardu MODS**. Návod je dostupný v kapitole 11.1.2 této metodiky (platí u monografií; ostatní typy dokumentů musí specifikovat registrátor při zahájení činnosti).

Jako standardní nastavení je touto metodikou doporučeno využívat tyto typy dokumentů:

MONOGRAPH, **MONOGRAPH_VOLUME**, **PERIODICAL_ISSUE**, nebo **OTHER**. Využití zbylých typů je dobrovolné.

11.1 Pokyny pro registrace jednotlivých typů dokumentů

Pokud je možné hodnotu nepovinného údaje z katalogizačního záznamu zjistit, pak je registrátor povinen ji dodat. Kromě vlastních popisných metadat, která se vážou k intelektuální entitě, je potřeba zaslat údaje k samotnému identifikovanému dokumentu. Tyto údaje podrobně specifikuje kapitola 11.1.6.

11.1.1 MONOGRAPH

Element se používá pouze pro jednosvazkové monografie. Na stejné šabloně je založen i typ THESIS určený pro vysokoškolské kvalifikační práce a typ OTHER s výše danými pravidly.

- title – titulní název, povinný údaj, pole 245\$a MARC 21 záznamu
- subTitle – další název, nepovinný údaj, pole 245\$b MARC 21 záznamu
- ccnb – číslo národní bibliografie, nepovinný údaj, pole 015\$a MARC 21 záznamu
- isbn – isbn monografie, nepovinný údaj, pole 020\$a MARC 21 záznamu
- otherId – jiný identifikátor, nepovinný údaj, pole 024\$a MARC 21 záznamu
- documentType – druh dokumentu, nepovinný údaj, pole LDR/007, 655\$a MARC 21 záznamu
- digitalBorn – původ monografie, pro monografie vzniklé digitalizací vždy false
- primaryOriginator – (hlavní) původce, nepovinný údaj, povolené hodnoty atributů jsou:
 - type="CORPORATION" – korporace, nepovinný údaj, pole 110\$a, \$b MARC 21 záznamu
 - type="AUTHOR" - Autor, nepovinný údaj, pole 100\$a, \$b, \$c, \$d MARC 21 záznamu
 - type="EVENT" - akce, nepovinný údaj, pole 111\$a, \$n, \$d, \$c MARC 21 záznamu
- otherOriginator – další původce, nepovinný údaj, pole 700\$a, \$b, \$c, \$d MARC 21 záznamu
- publisher – název vydavatele, nepovinný údaj, pole 260\$b MARC 21 záznamu
- place – místo vydání, nepovinný údaj, pole 260\$a MARC 21 záznamu
- year – rok vydání, nepovinný údaj, pole 260\$c MARC21záznamu

Metadatový popis pro nástroj resolver:

Následující části popisují vnitřní metadatový popis jednotlivých typů dokumentů. Popis říká, odkud je vhodné brát hodnoty pro plnění polí v databázi systému ČIDLO.

Tabulka 3: Doporučení na získání hodnoty „číslo ČNB“ z metadat

Název pole	Název podpole	Název pole v databázi
čČNB	čČNB	ccnb
Název pole	Šablona metadat	Povinnost
čČNB	<xs:element name="ccnb" maxOccurs="1" minOccurs="0" type="r:ccnbType"/>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	015 \$a	
DC kvalifikovaný	<dc:identifier type="uri">	
DC nekvalifikovaný	<dc:identifier>	
MODS	<mods:identifier type="uri">	
NUŠL	<dc:identifier>	
e-VŠKP	<dc:identifier>	
Ostatní		

Tabulka 4: Doporučení na získání hodnoty „ISBN“ z metadat

Název pole	Název podpole	Název pole v databázi
ISBN	ISBN	isbn
Název pole	Šablona metadat	Povinnost
ISBN	<xs:element name="isbn" type="r:isbnType" minOccurs="0" maxOccurs="1" />	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	020 \$a	
DC kvalifikovaný	<dc:identifier type="urn">	
DC nekvalifikovaný	<dc:identifier>	
MODS	<mods:identifier type="isbn">	
NUŠL	<dc:identifier>	
e-VŠKP	--	
Ostatní		

Jiný identifikátor jako např. vnitřní identifikátory lokálníchází mohou spadnout do

tohoto pole, přičemž je ale důležitá specifikace, o jaký typ identifikátoru se jedná.

Tabulka 5: Doporučení na získání hodnoty „jiný identifikátor“ z metadat

Název pole	Název podpole	Název pole v databázi
Jiný identifikátor	Jiný identifikátor	otherID
Název pole	Šablona metadat	Povinnost
Jiný identifikátor	<code><xs:element name="otherId" type="r:otherIdType" minOccurs="0" maxOccurs="1" /></code>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	024 \$a	
M21 – NUŠL/CDS	024 \$a	
DC kvalifikovaný	<code><dc:identifier type=""></code>	
DC nekvalifikovaný	<code><dc:identifier></code>	
MODS	<code><mods:identifier type=""></code>	
NUŠL	<code><dc:identifier></code>	
e-VŠKP	<code><dc:identifier></code>	
Ostatní		

Tabulka 6: Doporučení na získání hodnoty „druh dokumentu“ z metadat

Název pole	Název podpole	Název pole v databázi
Druh dokumentu	Druh dokumentu	documentType
Název pole	Šablona metadat	Povinnost
Druh dokumentu	<code><xs:element name="documentType" type="r:documentTypeType" minOccurs="0" maxOccurs="1" /></code>	Ne/ Ano u <i>entityType</i> OTHER.

Mapování		
Standard	Pole (Tag)	Poznámky
Aleph	druh dokumentu (BK)	
MARC21	LDR/007 – kód „m“, 655	
M21 – NUŠL/CDS	655 \$a	
DC kvalifikovaný	<code><dc:type></code>	
DC nekvalifikovaný	<code><dc:type></code>	
MODS	<code><mods:typeOfResource></code> ; <code><mods:genre></code>	
NUŠL	<code>dc:type; nusl:typeType</code>	

e-VŠKP	dc:type ; evskp:typeType; 655	
Ostatní		

Tabulka 7: Doporučení na získání hodnoty „hlavní název“ z metadat

Název pole	Název podpole	Název pole v databázi
Titulní název	Hlavní název	title
Název pole	Šablona metadat	Povinnost
Titulní název	<code><xs:element name="title" type="r:titleType" minOccurs="1" maxOccurs="1" /></code>	Ano

Mapování		
Standard	Pole (Tag)	Poznámky
Aleph		
MARC21	245 \$a	
M21 NUŠL/CDS	245 \$a	
DC kvalifikovaný	<dc:title>	
DC nekvalifikovaný	<dc:title>	V poli dc title se spojuje title a subTitle ve formátu „title: subTitle“
MODS	<mods:titleInfo><mods:title>	
NUŠL	<dc:title>	
e-VŠKP	<dc:title>	
Ostatní		

Tabulka 8: Doporučení na získání hodnoty „autor“ z metadat

Název pole	Název podpole	Název pole v databázi
Autor	Autor	primaryOriginator/Author
Název pole	Šablona metadat	Povinnost
Autor	<code><xs:element ref="r:primaryOriginator" minOccurs="0" maxOccurs="1" /> <xs:simpleType name="originatorTypeType"> <xs:enumeration value="AUTHOR" /></code>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	100 \$a, 100 \$b, 100 \$c, 100 \$d	
M21 NUŠL/CDS	720 \$a	
DC kvalifikovaný	<dc:creator>	
DC nekvalifikovaný	<dc:creator>	
MODS	<mods:name><mods:namePart>	
NUŠL	<dc:creator>	
e-VŠKP	<dc:creator>	
Ostatní		

Tabulka 9: Doporučení na získání hodnoty „korporace“ z metadat

Název pole	Název podpole	Název pole v databázi
Korporace	Korporace	primaryOriginator/Corporation
Název pole	Šablona metadat	Povinnost
Korporace	<xs:simpleType name="originatorTypeType"> <xs:enumeration value="CORPORATION" />	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	110 \$a, 110 \$b	
M21 NUŠL/CDS	720 \$a	
DC kvalifikovaný	<dc:creator>	
DC nekvalifikovaný	<dc:creator>	
MODS	<mods:name><mods:namePart>	
NUŠL	<dc:creator>	
e-VŠKP	----	
Ostatní		

Tabulka 10: Doporučení na získání hodnoty „akce“ z metadat

Název pole	Název podpole	Název pole v databázi
Akce	Akce	primaryOriginator/Event
Název pole	Šablona metadat	Povinnost
Akce	<xs:simpleType name="originatorTypeType"> <xs:enumeration value="EVENT" />	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	111 \$a, 111 \$n, 111 \$c, 111 \$d	
M21 NUŠL/CDS	711 \$a	
DC kvalifikovaný	---	
DC nekvalifikovaný	---	
MODS	<mods:name><mods:namePart>	
NUŠL	nusl:event ; nusl:eventPlace ; nusl:eventDates	
e-VŠKP	----	
Ostatní		

Tabulka 11: Doporučení na získání hodnoty „další původce“ z metadat

Název pole	Název podpole	Název pole v databázi
Další původce	Další původce	otherOriginator
Název pole	Šablona metadat	Povinnost
Další původce	<xs:element name="otherOriginator" type="r:otherOriginatorType" minOccurs="0" maxOccurs="1" />	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	700 \$a, 700 \$b, 700 \$c, 700 \$d	
M21 NUŠL/CDS	720 \$i	
DC kvalifikovaný	<dc:contributor>	
DC nekvalifikovaný	<dc:contributor>	
MODS	<mods:name><mods:namePart>	
NUŠL	<dc:contributor>	
e-VŠKP	----	
Ostatní		

Tabulka 12: Doporučení na získání hodnoty „místo vydání“ z metadat

Název pole	Název podpole	Název pole v databázi
Místo vydání	Místo vydání	place
Název pole	Šablona metadat	Povinnost
Místo vydání	<xs:element name="place" type="r:placeType" minOccurs="0" maxOccurs="1" />	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	260 \$a	
DC kvalifikovaný	<dc:publisher>	
DC nekvalifikovaný	<dc:publisher>	
MODS	<mods:originInfo><mods:place>	
NUŠL	----	
e-VŠKP	----	
Ostatní		

Tabulka 13: Doporučení na získání hodnoty „název nakladatele“ z metadat

Název pole	Název podpole	Název pole v databázi
Název nakladatele	Název nakladatele	publisher
Název pole	Šablona metadat	Povinnost
Název nakladatele	<pre><xs:element name="publisher" type="r:publisherType" minOccurs="0" maxOccurs="1" /></pre>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	260 \$b	
M21 NUŠL/CDS	260 \$b	
DC kvalifikovaný	<dc:publisher>	
DC nekvalifikovaný	<dc:publisher>	
MODS	<mods:originInfo><mods:publisher>	
NUŠL	<dc:publisher>	
e-VŠKP	----	
Ostatní		

Tabulka 14: Doporučení na získání hodnoty „rok vydání“ z metadat

Název pole	Název podpole	Název pole v databázi
Rok vydání	Rok vydání	year
Název pole	Šablona metadat	Povinnost
Rok vydání	<pre><xs:element name="year" type="xs:integer" minOccurs="0" maxOccurs="1" /></pre>	Ne

Mapování

Standard	Pole (Tag)	Poznámky
MARC21	260 \$c	
M21 NUŠL/CDS	046 \$k	
DC kvalifikovaný	<dcterms:issued> ; <dcterms:created> ; <dcterms:modified>	
DC nekvalifikovaný	<dc:date>	
MODS	<mods:originInfo><mods:dateIssued>	
NUŠL	<dcterms:dateCreated>	
e-VŠKP	<dcterms:created>	
Ostatní		

Vzorové namapování z metadat MODS pro tento typ, které metodiku doporučuje, vypadá takto (varianty se použijí v případě, že předchozí varianta není možná):

Tabulka 15: Vzorové mapování metadat MODS na pole databáze ČIDLO (monografie)

Název pole	1. varianta	2. varianta	3. varianta
title	//mods:titleInfo/mods:title (první hodnotu)		
subtitle	//mods:titleInfo/mods:subTitle (první hodnotu)		
ccnb	//mods:identifier[@type='ccnb']		
isbn	//mods:identifier[@type='isbn']		
otherId	//mods:identifier[@type='isrc']"		
documentType	z metsu pomocí metody getType()		
digitalBorn	pokaždé false		
primaryOriginator	//mods:name[@type='personal' and @usage='primary']//mods:namePart[not(@type='date')]	//mods:name[@type='corporate']	//mods:name[@type='conference']
otherOriginator	//mods:name[@type='personal' and not(@usage='primary')]/mods:namePart[not(@type='date')]		
publisher	//mods:originInfo/mods:publisher		
place	//mods:placeTerm[@type='text'][not(@authority)]		
year	//mods:originInfo/mods:dateIssued		

Zápis jména původce je závislý na podobě zápisu ve vstupních datech. Systém

ČIDLO akceptuje zápis ve tvaru *Božena Němcová i Němcová, Božena*. Do databáze však smí být odeslány jen vlastní jména původce, nikoli přidružená životní data. Platí pro všechny typy dokumentů.

11.1.2 MONOGRAPH_VOLUME

Element se používá pouze pro svazky monografie.

- monographTitle – titulní název celé monografie. Povinný údaj, pole 245 \$a MARC 21 záznamu
- volumeTitle – název svazku. Povinný údaj. Z metadat v digitalizaci (z PSP balíčku) nebo z pole 245 \$n MARC 21, pokud je k dispozici.
- ccnb – číslo národní bibliografie pro svazek monografie, pokud pro svazek monografie neexistuje, potom číslo národní bibliografie pro celou monografii, nepovinný údaj, pole 015 \$a MARC 21 záznamu
- isbn – isbn svazku monografie, nepovinný údaj, pole 020 \$a MARC 21 záznamu
- otherId – jiný identifikátor, nepovinný údaj, pole 024 \$a MARC 21 záznamu
- documentType – druh dokumentu, nepovinný údaj, pole LDR/007, 655 \$a MARC 21 záznamu
- digitalBorn – původ monografie, pro svazky monografie v digitalizaci NDK nebo digitalizované svazky monografií jiného původu bude vždy false; pro elektronicky vzniklé knihy bude mít hodnotu true
- primaryOriginator – (hlavní) původce, nepovinný údaj, povolené hodnoty atributů jsou:
 - type="CORPORATION" – korporace, nepovinný údaj, pole 110 \$a, \$b MARC 21 záznamu
 - type="AUTHOR" - Autor, nepovinný údaj, pole 100 \$a, \$b, \$c, \$d MARC 21 záznamu
 - type="EVENT" - akce, nepovinný údaj, pole 111 \$a, \$n, \$d, \$c MARC 21 záznamu
- otherOriginator – další původce, nepovinný údaj, pole 700 \$a, \$b, \$c, \$d MARC 21 záznamu
- publisher – název vydavatele, nepovinný údaj, pole 260 \$b MARC 21 záznamu
- place – místo vydání, nepovinný údaj, pole 260 \$a MARC 21 záznamu
- year – rok vydání, nepovinný údaj, pole 260 \$c MARC21 záznamu

Z hlediska podrobného metadatového popisu znamená tento typ jen přidání pole o názvu svazku. Pro zjištění, zda se jedná o svazek nebo ne, tedy zda je třeba použít masku tohoto typu dokumentu, je doporučen následující test při převodu popisu ve formátu MARC21 do standardu MODS:

```
<xsl:for-each select="marc:leader">
  <issuance>
    <xsl:choose>
      <xsl:when test="$leader7='a' or $leader7='c' or $leader7='d'"
        >monographic</xsl:when>
      <xsl:when test="$leader7='b' or $leader7='i' or $leader7='s'"
        >continuing</xsl:when>
      <xsl:when test="$leader7='m' and $leader19='a'"
        >multipart monograph</xsl:when>
      <xsl:when test="$leader7='m' and
../marc:datafield[@tag=245]/marc:subfield[@code='n' or @code='p']"
        >multipart monograph</xsl:when>
      <xsl:when test="$leader7='m' and $leader19=' '"
        >single unit</xsl:when>
      <xsl:otherwise>monographic</xsl:otherwise>
    </xsl:choose>
  </issuance>
</xsl:for-each>
```

Pokud je výsledkem transformace hodnota multipart monograph, pak je třeba při odesílání záznamu do resolveru šablonu dokumentu monographVolume.

Vzorové namapování z metadat MODS pro tento typ, které metodiku doporučuje, vypadá takto:

Tabulka 16: Vzorové mapování metadat MODS na pole databáze ČIDLO (svazek monografie)

Název pole	1. varianta	2. varianta	3. varianta
monograph Title	//mods:titleInfo/mods:title (první hodnotu)		
volumeTitle	//mods:titleInfo/mods:partNumber	//mods:titleInfo/mods:partName	
ccnb	//mods:identifier[@type='ccnb']		
isbn	//mods:identifier[@type='isbn']		
otherId	//mods:identifier[@type='isrc']"		
documentType	z METS pomoci metody getType()		

digitalBorn	False/true (pro digitalizované dokumenty vždy false)		
primaryOriginator	//mods:name[@type='personal' and @usage='primary']//mods:namePart[not(@type='date')]	//mods:name[@type='corporate']	//mods:name[@type='conference']
otherOriginator	//mods:name[@type='personal' and not(@usage='primary')]/mods:namePart[not(@type='date')]		
publisher	//mods:originInfo/mods:publisher		
place	//mods:placeTerm[@type='text'][not(@authority)]		
year	//mods:originInfo/mods:dateIssued		

11.1.3 PERIODICAL_ISSUE

Element se používá pouze pro čísla periodik.

- periodicalTitle – titulní název periodika, povinný údaj, pole 245 \$a MARC 21 záznamu
- volumeTitle – název ročníku (svazku) periodika. Povinný údaj, z metadat v digitalizaci (z PSP balíčku)
- issueTitle – název čísla periodika. Povinný údaj, z metadat v digitalizaci (z PSP balíčku)
- ccnb – číslo národní bibliografie, nepovinný údaj, pole 015 \$a MARC 21 záznamu
- issn – issn periodika, nepovinný údaj, pole 022 \$a MARC 21 záznamu
- otherId – jiný identifikátor, nepovinný údaj, pole 024 \$a MARC 21 záznamu
- documentType – druh dokumentu, pole LDR/007, 655 \$a MARC 21 záznamu
- digitalBorn – původ seriálu, pro digitalizaci v NDK nebo digitalizované seriály jiného původu bude vždy false
- primaryOriginator – (hlavní) původce, nepovinný údaj, povolené hodnoty atributů jsou:
 - type="CORPORATION" – korporace, nepovinný údaj, pole 110 \$a, \$b MARC 21 záznamu
 - type="AUTHOR" - Autor, nepovinný údaj, pole 100 \$a, \$b, \$c, \$d MARC 21 záznamu

type="EVENT" - akce, nepovinný údaj, pole 111 \$a, \$n, \$d, \$c MARC 21 záznamu

- otherOriginator- další původce, nepovinný údaj, pole 700 \$a, \$b, \$c, \$d MARC 21 záznamu
- publisher – název vydavatele, nepovinný údaj, pole 260 \$b MARC 21 záznamu
- place – místo vydání, nepovinný údaj, pole 260 \$a MARC 21 záznamu
- year – rok vydání, nepovinný údaj, pole 260 \$c MARC21 záznamu

Popis polí, která jsou odlišná od ostatních typů dokumentů:

Tabulka 17: Doporučení na získání hodnoty „druh dokumentu“ z metadat (číslo periodika)

Název pole	Název podpole	Název pole v databázi
Druh dokumentu	Druh dokumentu	documentType
Název pole	Šablona metadat	Povinnost
Druh dokumentu	<pre><xs:element name="documentType" type="r:documentTypeType" minOccurs="0" maxOccurs="1" /></pre>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
Aleph	druh dokumentu (SE)	
MARC21	LDR/007 – kód „s“, 655	
DC kvalifikovaný	<dc:type>	
DC nekvalifikovaný	<dc:type>	
MODS	<mods:typeOfResource> ; <mods:genre>	
NUŠL	---	
e-VŠKP	---	
Ostatní		
Ostatní		

Tabulka 18: Doporučení na získání hodnoty „ISSN“ z metadat (číslo periodika)

Název pole	Název podpole	Název pole v databázi
ISSN	ISSN	issn
Název pole	Šablona metadat	Povinnost
ISSN	<pre><xs:element name="issn" type="r:issnType" minOccurs="0" maxOccurs="1" /></pre>	Ne

Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	022 \$a	
DC kvalifikovaný	<dc:identifier type="urn">	
DC nekvalifikovaný	<dc:identifier>	
MODS	<mods:identifier type="issn">	
NUŠL	<dc:identifier>	
e-VŠKP	--	
Ostatní		

Doporučené mapování:

Tabulka 19: Vzorové mapování metadat MODS na pole databáze ČIDLO (číslo periodika)

Název pole	1. varianta	2. varianta	3. varianta
periodicalTitle	//mods:titleInfo/mods:title (první hodnotu)		
volumeTitle	//mods:titleInfo/mods:partNumber (z úrovně MODS_VOLUME)	//mods:titleInfo/mods:partName	
issueTitle	//mods:titleInfo/mods:partNumber (z úrovně MODS_ISSUE)	//mods:titleInfo/mods:partName	
ccnb	//mods:identifier[@type='ccnb']		
issn	//mods:identifier[@type='issn']		
otherId	//mods:identifier[@type='isrc']"		
documentType	z metsu pomocí metody getType()		
digitalBorn	False/true (pro digitalizované dokumenty vždy false)		
primaryOriginator	//mods:name[@type='personal' and @usage='primary']//mods:namePart[not(@type='date')]	//mods:name[@type='corporate']	//mods:name[@type='conference']
otherOriginator	//mods:name[@type='personal' and not(@usage='primary')]/mods:namePart[not(@type='date')]		
publisher	//mods:originInfo/mods:publisher		
place	//mods:placeTerm[@type='text'][not(@authority)]		
year	//mods:originInfo/mods:dateIssued		

Při nejasnostech jak vyplňovat vstupní údaje při popisu periodik v průběhu digitalizace, je nutné se řídit *Pravidly popisu* dostupnými z <http://www.ndk.cz/standards-digitalizace/metadata> [citováno 2014-09-25].

11.1.4 PERIODICAL_VOLUME

Element se používá pouze pro ročník periodik.

- periodicalTitle - název periodika, povinný údaj, pole 245 \$a MARC 21 záznamu
- volumeTitle – název ročníku, povinný údaj, z metadat v digitalizaci (z PSP balíčku)
- ccnb – číslo národní bibliografie, nepovinný údaj, pole 015 \$a MARC 21 záznamu
- issn – issn periodika, nepovinný údaj, pole 022 \$a MARC 21 záznamu
- otherId – jiný identifikátor, nepovinný údaj, pole 024 \$a MARC 21 záznamu
- documentType – druh dokumentu, pole LDR/007, 655 \$a MARC 21 záznamu
- digitalBorn – původ seriálu, pro digitalizaci v NDK pro novou digitalizaci v NDK nebo digitalizované seriály jiného původu bude vždy false
- primaryOriginator – (hlavní) původce, nepovinný údaj, povolené hodnoty atributů jsou:
 - type="CORPORATION" – korporace, nepovinný údaj, pole 110 \$a, \$b MARC 21 záznamu
 - type="AUTHOR" - Autor, nepovinný údaj, pole 100 \$a, \$b, \$c, \$d MARC 21 záznamu
 - type="EVENT" - akce, nepovinný údaj, pole 111 \$a, \$n, \$d, \$c MARC 21 záznamu
- OtherOriginator - další původce, nepovinný údaj, pole 700 \$a, \$b, \$c, \$d MARC 21 záznamu
- publisher – název vydavatele, nepovinný údaj, pole 260 \$b MARC 21 záznamu
- place – místo vydání, nepovinný údaj, pole 260 \$a MARC 21 záznamu
- year – rok vydání, nepovinný údaj, pole 260 \$c MARC21 záznamu

Popis je obdobný jako u předchozího typu.

11.1.5 ANALYTICAL

Element se používá pouze pro analytika.

- title – titulní název. Povinný údaj, pole 245 \$a MARC 21 záznamu
- subTitle – další název, nepovinný údaj, pole 245 \$b MARC 21 záznamu
- otherId – jiný identifikátor, nepovinný údaj, pole 024 \$a MARC 21 záznamu
- documentType – druh dokumentu, pole LDR/007, 655 \$a MARC 21 záznamu
- primaryOriginator – původce, nepovinný údaj, povolené hodnoty atributů jsou:
type="CORPORATION" – korporace, nepovinný údaj, pole 110 \$a, \$b MARC 21 záznamu
type="AUTHOR" - Autor, nepovinný údaj, pole 100 \$a, \$b, \$c, \$d MARC 21 záznamu
type="EVENT" - akce, nepovinný údaj, pole 111 \$a, \$n, \$d, \$c MARC 21 záznamu
- otherOriginator - další původce, nepovinný údaj, pole 700 \$a, \$b, \$c, \$d MARC 21 záznamu
- sourceDocument – informace o zdrojovém dokumentu analytika, povinná
- title – název zdrojového dokumentu (může být monografie nebo periodikum), pokud je k dispozici, povinný údaj
- volumeTitle – název ročníku (svazku) periodika (pokud je zdrojovým dokumentem periodikum), pokud je k dispozici, nepovinný údaj
- issueTitle – název čísla periodika (pokud je zdrojovým dokumentem periodikum), pokud je k dispozici, nepovinný údaj
- ccnb – ČČNB zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- isbn – ISBN zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- issn – ISSN zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- otherId – další identifikátor zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- publisher – název vydavatele zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- place – místo vydání zdrojového dokumentu, pokud je k dispozici, nepovinný údaj
- year – rok vydání zdrojového dokumentu, pokud je k dispozici, nepovinný údaj

Tabulka 20: Vzorové mapování metadat MODS na pole databáze ČIDLO (článek)

Název pole	1. varianta	2. varianta	3. varianta
title	//mods:titleInfo/mods:title		
subTitle	//mods:titleInfo/mods:subTitle		
otherId	//mods:identifier[@type='isrc']		
documentType	//mods:genre		
primaryOriginator	//mods:name[@type='personal' and @usage='primary']//mods:namePart[not(@type='date')]	//mods:name[@type='corporate']	
otherOriginator	//mods:name[@type='personal' and not(@usage='primary')] //mods:namePart[not(@type='date')]		
digitalBorn	False/true (pro digitalizované dokumenty vždy false)		

Zbylá pole identifikující zdrojový dokument se vyplňují podle šablony pro PERIODICAL_ISSUE.

11.1.6 DIGITAL DOCUMENT

Následující údaje se vztahují k vlastnímu digitálnímu dokumentu bez ohledu na jeho typ. Tedy na tu část, která je identifikovaná pomocí identifikátoru URN:NBN.

- **archvierId** – povinný údaj. V případě, že registrátor vystupuje zároveň jako archivátor, tedy sám dlouhodobě ukládá své dokumenty, se hodnota nevyplňuje. V případě, že dokumenty budou k dlouhodobému uložení svěřeny jiné instituci, bude vyplněn kód této instituce coby archivátora. Tento údaj se nezapisuje do metadat dokumentu (nezaměňovat s hodnotou Archivist), je zadán pouze v requestu pro resolver. Jako příklad lze uvést kód LTP úložiště NKČR = 45.
- **digitalDocument id** – povinný údaj, doplňuje samotný systém, slouží jako vnitřní identifikátor, přidělí se v okamžiku registrace, nezapisuje se do metadat.
- **urnNbn** – URN:NBN dokumentu, bude vyplněno pouze v případě, kdy nebyl resolver k dispozici a URN:NBN bylo přiděleno z dopředu rezervovaných identifikátorů.

V opačném případě zůstává prázdné a URN:NBN je přiděleno jako odpověď na tento regist.

- **registrarScopeIdentifiers** – seznam identifikátorů digitálního dokumentu, které jsou jedinečné v rámci registrátora
- **registrarScopeIdentifier** - jedinečný identifikátor

Hodnota identifikátoru má maximální délku 60 znaků. Může obsahovat číslice, znaky malá a velká písmena a dále tyto znaky: \$ - _ . + ! * ' () , :

V rámci registrátora nemohou mít dva různé digitální dokumenty stejnou hodnotu identifikátoru (pro stejný typ)

- **atribut type** – typ identifikátoru

typ identifikátoru musí mít do 20 znaků. Skládá se z číslic, znaků malých a velkých písmen a těchto dalších znaků: _ \ - :

financed – způsob financování, nepovinný údaj, př. IOP-NDK, norské fondy.

contractNumber – číslo zakázky, nepovinný údaj

Tabulka 21: Vzorové mapování metadat na pole databáze ČIDLO (digitální dokument)

Název pole	hodnota
archvierld	Viz popis
urnnbn	///mods:identifier[@type='urnnbn']
registrarScopeIdentifiers	vložen následující registrarScopeIdentifier
registrarScopeIdentifier_type	uuid
registrarScopeIdentifier_value	Hodnota cdmId
financed	
contractNumber	

To, co je obsaženo v TYPE i VALUE, je čistě na vůli registrátora, samotný resolver o hodnotách nic nepředpokládá, jen je ukládá do databáze a při resolvování hledá záznam s určeným typem a hodnotou. Přes REST API je možné hodnoty identifikátorů zjišťovat, měnit, odebírat (na základě typu).

V názvosloví této metodiky je pojem digitální dokument používán podle definice uvedené v kapitole 4. Pokud se zde mluví o údajích popisující Digital Document, tak se tímto pojmem rozumí digitální reprezentace a zároveň vlastní dokument, jenž je označen pomocí identifikátoru. Tato část popisu zachycuje unikátní údaje.

```

- <response xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance http://resolver.nkp.cz/api/v3/response.xsd">
- <digitalDocument id="493055">
- <urnNbn>
  <value>urn:nbn:cz:nk-000hg5</value>
</urnNbn>
  <financed>IOP-NDK</financed>
  <created>2013-09-29T14:46:23.521+02:00</created>
- <intellectualEntity type="MONOGRAPH">
  <created>2013-09-29T14:46:23.473+02:00</created>
- <titleInfo>
  <title>Mezidruhové křížení rodu Brassica</title>
</titleInfo>
  <ccnb>cnb000061787</ccnb>
  <isbn>8021301120</isbn>
  <digitalBorn>>false</digitalBorn>
  <primaryOriginator type="AUTHOR">Bechyně, Miroslav</primaryOriginator>
- <publication>
  <place>Praha</place>
  <year>1991</year>
</publication>
</intellectualEntity>
- <technicalMetadata>
  <format version="1.0">image/jp2</format>
  <extent>70 x JPEG2000</extent>
- <resolution>
  <horizontal>300</horizontal>
  <vertical>300</vertical>
</resolution>
  <compression>JPEG2000</compression>
- <color>
  <model>RGB</model>
  <depth>8</depth>
</color>
- <pictureSize>
  <width>1654</width>
  <height>2410</height>
</pictureSize>
</technicalMetadata>
- <registrarScopeIdentifiers>
  <id type="uuid">70ac8860-d400-11e2-8f90-005056827e51</id>
</registrarScopeIdentifiers>
- <registrar code="nk" id="73">
  <name>Národní knihovna České republiky</name>
  <created>2013-01-09T15:40:37.740+01:00</created>
  <modified>2014-06-09T13:13:34.088+02:00</modified>
- <registrationModes>
  <mode name="BY_RESOLVER" enabled="true"/>
  <mode name="BY_REGISTRAR" enabled="false"/>
  <mode name="BY_RESERVATION" enabled="true"/>
</registrationModes>
</registrar>
- <archiver id="45">
  <name>Národní knihovna České republiky</name>
  <description>nk</description>
  <created>2012-04-18T15:21:37.227+02:00</created>
  ...

```

Obrázek č. 4 – XML soubor zobrazující správně vytvořený záznam intelektuální entity a digitálního dokumentu bez údajů o digitální instanci.

11.2 Technická metadata zasílaná systému ČIDLO

Do databáze je možné vkládat i technická metadata, která se musí vztahovat k prvnímu obrazovému souboru v AIP balíčku. Je možné vyplňovat tato pole:

- **extent** – rozsah, nepovinný údaj, označuje počet počítačových souborů (reprezentujících obrazová data) archivního balíčku, např. 150xjpeg2000, váže se pouze k archivnímu balíčku jako celku
- **format** – formát (např. JPEG) digitálního dokumentu, povinný údaj, bude vyplněno pro dokumenty procházející digitalizací, váže se pouze k archivnímu balíčku jako celku, nebo k prvnímu počítačovému souboru tohoto archivního balíčku
- **atribut: version:** verze formátu, povinný údaj
- **resolution** – rozlišení nastavené na scanneru v dpi, nepovinný údaj
- **horizontal** – šířka, např. 300
- **vertical** – výška, např. 300
- **compression** – komprese, nepovinný údaj
- **atribut: ratio** – kompresní poměr, nepovinný údaj
- **color** – informace o barvě, nepovinný údaj
- **model** – barevný model, např. RGB, nepovinný údaj
- **depth** – barevná hloubka v bitech, např. 24, povinný údaj
- **iccProfile** – nepovinný údaj
- **pictureSize** – reálná velikost prvního z obrazů archivní kopie v pixelech, nepovinný údaj,
- **width** – šířka, např. 800
- **height** – výška, např. 600

Tabulka 22: Doporučené mapování z metadat ve standardu MIX

format	value	//mix/BasicDigitalObjectInformation/FormatDesignation/formatName
	version	//mix/BasicDigitalObjectInformation/FormatDesignation/formatVersion
resolution	horizontal	výpočet numerator/denominator (numerator = //mix/ImageAssessmentMetadata/SpatialMetrics/XSamplingFrequency/Numerator, denominator = //mix/ImageAssessmentMetadata/SpatialMetrics/XSamplingFrequency/Denominator)

	vertical	výpočet numerator/denominator (numerator = //mix/ImageAssessmentMetadata/SpatialMetrics/YSamplingFrequency/ Numerator, denumenatro = //mix/ImageAssessmentMetadata/SpatialMetrics/YSamplingFrequency/ Denominator)
compression	value	//mix/BasicDigitalObjectInformation/Compression/CompressionS cheme
	ratio	//mix/BasicDigitalObjectInformation/Compression[0]/Compressio nRatio
color	model	//mix/BasicImageInformation/BasicImageCharacteristics/Photome tricInterpretation/colorSpace
	depth	//mix/ImageAssessmentMetadata/ImageColorEncoding/BitsPerSa mple/bitsPerSampleValue
iccProfile		//mix/BasicImageInformation/BasicImageCharacteristics/Photome tricInterpretation/ColorProfile/IccProfile
pictureSize	Height	(26 * height_of_image_in_px) / vertical_dpi;
	Width	(26 * width_of_image_in_px) / horizontal_dpi;

Příklady technických metadat ve formátu MIX, ze kterých mají být získávány údaje pro vyplnění polí databáze systému ČIDLO.

Formát, verze formátu:

```
<mix:BasicDigitalObjectInformation>
<mix:FormatDesignation>
<mix:formatName>image/jp2</mix:formatName>
<mix:formatVersion>1.0</mix:formatVersion>
</mix:FormatDesignation>
```

Rozlišení:

```
<mix:samplingFrequencyUnit>in.</mix:samplingFrequencyUnit>
<mix:xSamplingFrequency>
<mix:numerator>300000</mix:numerator>
<mix:denominator>1000</mix:denominator>
</mix:xSamplingFrequency>
<mix:ySamplingFrequency>
<mix:numerator>300000</mix:numerator>
<mix:denominator>1000</mix:denominator>
</mix:ySamplingFrequency>
```

Kompresní algoritmus:

```
<mix:Compression>
<mix:compressionScheme>JPEG2000</mix:compressionScheme>
```

</mix:Compression>

Barevný model a barevná hloubka:

<mix:PhotometricInterpretation>

<mix:colorSpace>RGB</mix:colorSpace>

</mix:PhotometricInterpr<mix:ImageColorEncoding>

<mix:BitsPerSample>

<mix:bitsPerSampleValue>8</mix:bitsPerSampleValue>

<mix:bitsPerSampleValue>8</mix:bitsPerSampleValue>

<mix:bitsPerSampleValue>8</mix:bitsPerSampleValue>

<mix:bitsPerSampleUnit>integer</mix:bitsPerSampleUnit>

</mix:BitsPerSample>

<mix:samplesPerPixel>3</mix:samplesPerPixel>

</mix:ImageColorEncoding>

Velikost:

<mix:BasicImageCharacteristics>

<mix:imageWidth>1719</mix:imageWidth>

<mix:imageHeight>2664</mix:imageHeight>

</mix:BasicImageCharacteristics>

12. RESOLVOVACÍ FUNKCE SYSTÉMU ČIDLO

Resolvování nebo-li přesměrování znamená reakci resolveru na URL (v domluveném formátu) obsahující URN:NBN identifikátor. Na základě dostupných dat se resolver zachová jedním z těchto tří způsobů:

1. Odpoví kódem 302 a přesměruje tak prohlížeč do některé digitální knihovny na URL konkrétní digitální instanci požadovaného digitálního dokumentu.
2. Odpoví kódem 302 a přesměruje do webové části systému (modul *web*), kde se zobrazí metadata v lidsky čitelné podobě.
3. Zobrazí metadatový záznam dokumentu v xml.

Přesné chování resolveru je možné vynutit parametry *action* a *format*.

Bez upřesňujících parametrů se pokusí přesměrovat do digitální knihovny. Pokud to není možné (neexistuje žádná digitální instance identifikovaného dokumentu), zobrazí se lidsky čitelný metadatový záznam ve webové části systému.

Uživatel znající hodnotu identifikátoru může získat přístup k žádanému dokumentu třemi způsoby, pokud bude využívat identifikátor a možnosti, které nabízí systém ČIDLO a spolupracující systémy.

- 1) Vyhledání pomocí webového rozhraní systému ČIDLO. V záložce Vyhledávání uživatel zadá identifikátor URN:NBN a dostane záznam v databázi, který obsahuje bibliografický popis i technická metadata původního archivního dokumentu a odkaz do digitální knihovny. Kromě vyhledávání podle identifikátoru může ve webovém rozhraní hledat i pomocí identifikátoru čČNB (číslo České národní bibliografie), čísel ISSN, ISBN a názvových údajů (např. titul monografie, název periodika) nebo jejich částí.
- 2) K dokumentu se může dostat i přímo použitím směrování resolveru a to tak, že do adresního řádku vyhledávače zadá <https://resolver.nkp.cz/urn:nbn:cz:mzk-000bf5>. Resolver ho automaticky přesměruje na příslušnou digitální instanci. Je-li k dokumentu připojeno více instancí, vybere tu první v pořadí a v případě nedostupnosti zkouší i další.
- 3) V případě přístupu přes OPAC nabízí resolver směrování na základě zdroje přístupu. Tedy pokud uživatel klikne na odkaz v katalogu (pole 856), bude přesměrován primárně na digitální instanci náležející do knihovny provozované vlastníkem katalogu. Např. když budou v katalogu *Aleph* Moravské zemské knihovny v záznamech odkazy typu <http://resolver.nkp.cz/urn:nbn> a návštěvník katalogu na tento

odkaz klikne, potom resolver nalezne v hlavičce řetězec, který ho bude informovat o zdroji dotazu a nejdříve ověří, zda k danému katalogu je přiřazena i digitální knihovna, která obsahuje digitální instanci požadovanou uživatelem.

13. HROMADNÉ SKLÍZENÍ POMOCÍ OAI-PMH

Rámec OAI-PMH je určen pro hromadné sklizení metadatových záznamů. Definuje aplikační rozhraní serveru a komunikační protokol, kterým je sklizení realizováno. Strana nabízející metadata se nazývá *Provider* a klient, který komunikaci inicializuje a metadata sklízí, je označován jako *Harvester*. Modul OAI-PMH mohou ve svém jmenném prostoru spravovat samotní registrátoři nebo po dohodě může jejich správu převzít centrální autorita. Podrobnější návod pro technické nastavení OAI protokolu lze nalézt na projektové stránce systému ČIDLO <https://code.google.com/p/czidlo/wiki/Introduction?tm=6>

Funkci OAI-PMH Data provider realizuje aplikace **OaiPmhProvider**. Příkladem pro Resolver dostupný na <http://resolver.nkp.cz> je provider dostupný na <http://resolver.nkp.cz/OaiPmhProvider/provider>, přičemž existuje také jednoduché webové rozhraní - <http://resolver.nkp.cz/OaiPmhProvider>.

Operace protokolu OAI-PMH

Protokol definuje těchto 6 operací:

- **Identify** zobrazuje základní informace o implementaci Provideru a konkrétní možnosti implementace, jako je granularita časových známek nebo verze protokolu.
- **ListMetadataFormats** vrací seznam metadatových formátů, které Provider podporuje. Pokud je definována hodnota parametru *identifier*, jsou zobrazeny jen formáty, které jsou dostupné pro identifikovaný záznam, pokud tento existuje.
- **GetRecord** vrací záznam vybrané položky ve formátu určeném parametrem *metadataPrefix*. Záznam musí být identifikován parametrem *identifier*.
- **ListSets** vrací seznam množin. Množiny jsou pojmenované skupiny záznamu. Množiny se mohou překrývat, mohou být budovány hierarchicky a vznikat tak podmnožiny.
- **ListRecords** vrací seznam záznamu ve formátu určeném parametrem *metadataPrefix*. Výběr záznamu je možné dále upřesnit nepovinnými parametry *from*, *until* a *set*.
- **ListIdentifiers** vrací seznam identifikátorů v určeném metadatovém formátu. Vrací se jen hlavičky, nikoliv celé metadatové záznamy. Stejně jako u operace ListRecords lze i zde výsledný seznam ovlivnit parametry *from*, *until* a *set*.

Je-li potřeba sklizeň omezit například jen na některou kolekci záznamů, je možné použít mechanismus množin. Pokud Provider definuje množinu, může Harvester její jméno použít jako hodnotu parametru *set* a omezit tak sklizené výsledky na obsah množiny. Dalším

mechanismem selektivní sklizně jsou časová razítka. Ta jsou obsahem parametru *from* a *until*, je-li potřeba omezit výsledky na záznamy vytvořené nebo změněné v určitém časovém intervalu. Jedná se o uzavřený interval, ve kterém *from* určuje počátek a *until* konec. Pokud není parametr *from* definován, je za počátek intervalu považováno záporné nekonečno, resp. kladné nekonečno v případě *until*. Oba mechanismy je možné kombinovat a sklízet tak jen z určeného časového rozsahu a zároveň vybrané množiny.

Sety

Pro každého registrátora je definován set se jménem **registrar:** + *kód registrátora*, např. **registrar:aba008**. Je tedy možné sklídit jen záznamy vlastněné vybraným registrátorem.

13.1 OAI Adapter

Dále systém ČIDLO disponuje nástrojem OAI Adapter. OAI Adapter je nástroj umožňující registrátorům vkládat a udržovat aktuální odkazy na zveřejněné instance digitálních dokumentů s přiřazeným URN:NBN identifikátorem. Je určen pro ty digitální knihovny, které disponují OAI-PMH rozhraním. Skládá se z těchto komponent:

- OAI Harvester - sklízí záznamy z OAI-PMH rozhraní digitální knihovny
- Transformer - transformuje sklizené záznamy na data pro *CZIDLO API*. Tedy na vstup operace *registrace digitálního dokumentu* a na vstup operace *import digitální instance*. Transformace jsou realizovány pomocí technologie XSLT (verze 1.0). Jsou proto potřeba vždy dva XSLT soubory.
- Resolver connector - modul přistupující k *CZIDLO API*. Provádí operace *registrace digitálního dokumentu*, *import digitální instance* a některé další.

OAI Adapter tedy primárně zajišťuje aktuálnost dig. instance. To je potřeba například v následujících situacích:

- Změní se url digitální knihovny a proto i url na jednotlivé tituly (url je hlavní částí dig. instance).
- Změní se viditelnost některých záznamů.
- Registrátor nemá nástroje pro registraci dokumentu. Pokud má jeho digitální knihovna OAI-PMH provider, může registrátor použít tento nástroj také k registraci dig. dokumentu, nejen k importu dig. instance.

OAI Adapter prochází všechny záznamy a aktualizuje jen ty, které se změnilly.

14. PRAVIDLA PRO ŽIVOTNÍ CYKLUS DOKUMENTU A IDENTIFIKÁTORU

Jeden digitální dokument, který vznikl digitalizací tištěné předlohy a jemuž byl v rámci fáze finalizace produkčního balíčku přidělen URN:NBN identifikátor, následně existuje ve dvou verzích – jako archivní dokument a jako uživatelský dokument. Obě verze získávají tentýž URN:NBN identifikátor. Ten musí být vepsán jak do metadat balíčku AIP, tak do metadat balíčku DIP (respektive do balíčků SIP1 a SIP2, pokud archivní a uživatelský dokumenty vznikají souběžně během digitalizace). Ani jeden z nich pak již nesmí být měněn. Popisná metadata uživatelské verze musí odpovídat popisným metadatům archivního balíčku v poměru 1:1. Všechny verze digitálního dokumentu musejí být odvozeny buď z původního archivního dokumentu nebo původního uživatelského dokumentu, pokud archivní a uživatelský dokument vznikaly souběžně při digitalizaci.

V případě budoucích formátových migrací uskutečněných v digitálním repozitáři s digitálním dokumentem⁸ platí, že nové verze vzniklé těmito migracemi dědí identifikátor URN:NBN po původním archivním dokumentu. Tyto migrace musejí být provedeny tak, aby neporušily autenticitu digitálního dokumentu, tj. podle doporučení nejnovějších verzí norem ISO 16363 a ISO 14721 a dalších osvědčených postupů užívaných v knihovnické komunitě. Technická identifikátorová metadata uložená v databázi resolveru nejsou při formátové migraci aktualizována a odpovídají nadále původnímu archivnímu dokumentu.

Pokud v rámci životního cyklu digitálního dokumentu dojde k jeho významné změně (v rámci jednoho vlastníka), musí si registrátor nechat přidělit nový identifikátor a vytvořit vazbu na následnický dokument. Dojde-li ke změně velkého rozsahu, musí registrátor deaktivovat záznam digitálního dokumentu (**kód 200 pro automatizované postupy pomocí API nebo manuálně pomocí webového rozhraní u jednotlivých záznamů**) a registrovat dokument s novým identifikátorem URN:NBN bez vytvoření následnické vazby. Pokud registrátor trvale smazal digitalizovaný dokument nebo ho trvale znepřístupnil ve všech instancích, pak musí deaktivovat záznam dokumentu v systému ČIDLO a to deaktivací bez následníka.

U nesignifikantních změn má registrátor možnost měnit záznam v databázi, aniž by bylo třeba přidělit nový identifikátor. V případě deaktivace dokumentu je navíc potřeba provést samostatné vymazání dalších identifikátorů (typicky uuid), pokud registrátor chce tyto použít v další činnosti (resolver provádí kontrolu unikátnosti i dalších identifikátorů).

⁸ Například migrace souborů v JPEG 2000 do nějakého budoucího nového grafického formátu.

Vymazání se provádí operací DELETE.

Např. DELETE <http://resolver.nkp.cz/api/v3/resolver/urn:nbn:cz:nk-000005/registrarscopeidentifiers/uuid>
Nebo příkaz <http://resolver.nkp.cz/api/v3/resolver/urn:nbn:cz:boa001-000005/registrarscopeidentifiers> pro odstranění všech identifikátorů.

V odpovědi je zaslána hodnota všech odstraněných identifikátorů.

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <registrarscopeidentifiers>
    <id type="sysno">NKP01:000721328</id>
    <id type="uuid">0c6f5a60-db01-11e1-b015-00505682629d</id>
  </registrarscopeidentifiers>
</response>
```

V některých případech postačí jen změna konkrétního identifikátoru, která nevyžaduje přidělení nového URN:NBN identifikátoru.

Request (slouží pro vložení a aktualizaci identifikátoru)

```
PUT http://resolver.nkp.cz/api/v3/resolver/urn:nbn:cz:boa001-000005/registrarscopeidentifiers/TEST2
```

Tělo dotazu pak obsahuje novou hodnotu identifikátoru, zde: 1234

Response

Pokud digitální dokument identifikátor tohoto typu zatím neměl (vložení)

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <id type="TEST2">1234</id>
</response>
```

Pokud identifikátor existoval, je jeho hodnota změněna (aktualizace)

```
<?xml version="1.0" encoding="UTF-8"?>
<response xmlns="http://resolver.nkp.cz/v3/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-instance
http://resolver.nkp.cz/api/v3/response.xsd">
  <id type="TEST2" previousValue="12345">1234</id>
</response>
```

Registrátor nebo správce digitálního repozitáře se musí při správě dokumentů identifikovaných pomocí URN:NBN identifikátoru řídit podle představených pravidel.

Především zajistit, aby všechny verze původního archivního i uživatelského dokumentu dědily tentýž identifikátor a pokud dojde ke změně dokumentu, musí provést deaktivaci původního záznamu a jemu přiděleného identifikátoru. Dále, je-li to možné, musí zajistit vytvoření vazby "následník – předchůdce" v systému ČIDLO.

Změny, ke kterým může docházet, jsou tyto:

1. Změny vyžadující deaktivaci původního záznamu dokumentu a následnou novou registraci bez vazby na původní dokument:

- Zjištění skutečnosti, že vlastníkem dokumentu je jiná instituce, než kterou reprezentuje daný registrátor.
- Zjištění, že dokument byl registrován pod chybným typem dokumentu.
- Zjištění, že hlavní identifikační údaje jako název a autor byly chybné (nejde o překlepy, ale o záměny jmen – např. místo knihy s názvem Babička se dokument ve skutečnosti jmenuje Honzíkova cesta; pokud by chyba byla jen v tom, že v původním popisu je místo Babička uvedeno Babičk, lze postupovat podle bodu 2).

2. Změny vyžadující deaktivaci původního záznamu dokumentu a vytvoření vazby na následnický dokument:

- Sloučení dokumentů – sloučením dokumentů se rozumí stav, kdy se dvě a více odlišných intelektuálních entit spojí do jedné intelektuální entity.
- Rozdělení dokumentu – rozdělení jedné intelektuální entity do dvou nebo více intelektuálních entit.
- Oprava popisu dokumentu, která nepovede k záměně intelektuálních entit (oprava překlepů nebo chybějící části názvu).
- Signifikantně jiná digitalizace téže intelektuální entity v rámci jednoho vlastníka.
- Doskenování nebo přeskenování původního dokumentu v rámci stejné digitalizace (tedy doplnění nebo vyměnění obrazových souborů náležející k popisným metadatům – v takovém případě musí být důvod vytvoření vazby na následníka zapsán v poznámce [note] requestu do resolveru).

3. Změny nevyžadující deaktivaci původního dokumentu:

- Doplnění dosud chybějících údajů (např. identifikátor čČNB)
- Změna v údajích popisujících digitální instanci (včetně změn identifikátorů jako např.

uuid)

Pokud je potřeba přeregistrovat digitální dokument a přiřadit mu nový identifikátor URN:NBN, spojit několik digitálních dokumentů v jeden anebo naopak rozdělit digitální dokument v několik nových, je vhodné použít systém následníků a předchůdců. Pokud je digitální dokument v roli předchůdce jednoho nebo více digitálních dokumentů, je předchůdce vždy deaktivován (resp. jeho URN:NBN). Následník je aktivní, pokud není zároveň předchůdcem jiného digitálního dokumentu. Vztah *předchůdce-následník* má kardinalitu M:N, předchůdce tedy může mít jednoho a více následníků a následník může mít jednoho nebo více předchůdců. Definice této vazby je součástí operace *registrace digitálního dokumentu* a je realizována záznamem v metadatech nově registrovaného digitálního dokumentu. U všech příkladů se vyskytuje element `urnNbn/value`, jedná se tedy o mód `BY_REGISTRAR` nebo `BY_RESERVATION`. Slučovat a rozpojovat lze pouze dokumenty stejných typů. Mění-li se typ dokumentu, nelze použít vazbu předchůdce – následník, ale je třeba provést deaktivaci bez následníka a nově vzniklé dokumenty registrovat znovu.

14.1 Návod na postupy při slučování a rozpojování dokumentů

Spojení dvou digitálních dokumentů v jeden

```
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  ...
  <r:digitalDocument>
    ...
    <r:urnNbn>
      <r:value>urn:nbn:cz:boa001-00003b</r:value>
      <r:predecessor value="urn:nbn:cz:boa001-000030" note="spojení
dvou DD v jeden"/>
      <r:predecessor value="urn:nbn:cz:boa001-00003a" note="spojení
dvou DD v jeden"/>
    </r:urnNbn>
    ...
  </r:digitalDocument>
</r:import>
```

Vzniknou dvě vazby *předchůdce-následník*:

- `urn:nbn:cz:boa001-000030` -> `urn:nbn:cz:boa001-00003b`
- `urn:nbn:cz:boa001-00003a` -> `urn:nbn:cz:boa001-00003b`

Oba předchůdci (`urn:nbn:cz:boa001-000030`, `urn:nbn:cz:boa001-00003a`) jsou deaktivováni, pokud již deaktivováni nebyli. Deaktivované URN:NBN při resolvování klientův webový prohlížeč nikdy nepřesměruje do digitální knihovny, ale vždy zobrazí záznam v CZIDLO, jehož součástí jsou URN:NBN následníků.

Nahrazení digitálního dokumentu (např. změna v metadatech)

```
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  ...
  <r:digitalDocument>
    ...
    <r:urnNbn>
      <r:value>urn:nbn:cz:boa001-00003b</r:value>
      <r:predecessor value="urn:nbn:cz:boa001-00003a" note="zásadní
chyby v metadatech"/>
    </r:urnNbn>
    ...
  </r:digitalDocument>
</r:import>
```

Vznikne jedna vazba *předchůdce-následník*, konkrétně urn:nbn:cz:boa001-00003a -> urn:nbn:cz:boa001-00003b, digitální dokument *urn:nbn:cz:boa001-00003a* je deaktivován.

Rozdělení digitálního dokumentu ve dva

První digitální dokument:

```
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  ...
  <r:digitalDocument>
    ...
    <r:urnNbn>
      <r:value>urn:nbn:cz:boa001-00003b</r:value>
      <r:predecessor value="urn:nbn:cz:boa001-00003a" note="rozdělení
DD na dva"/>
    </r:urnNbn>
    ...
  </r:digitalDocument>
</r:import>
```

Druhý digitální dokument:

```
<r:import xmlns:r="http://resolver.nkp.cz/v3/">
  ...
  <r:digitalDocument>
    ...
    <r:urnNbn>
      <r:value>urn:nbn:cz:boa001-00003c</r:value>
      <r:predecessor value="urn:nbn:cz:boa001-00003a" note="rozdělení
DD na dva"/>
    </r:urnNbn>
    ...
  </r:digitalDocument>
</r:import>
```

Vzniknou dvě vazby *předchůdce-následník*:

- urn:nbn:cz:boa001-00003a -> urn:nbn:cz:boa001-00003b
- urn:nbn:cz:boa001-00003a -> urn:nbn:cz:boa001-00003c

Digitální dokument *urn:nbn:cz:boa001-00003a* je deaktivován.

O AUTORECH

Hlavními autory této metodiky jsou Zdeněk Vašek a Ladislav Cubr (autor koncepce systému). Na vzniku textů, které jsou částečně nebo zcela součástí metodiky, externě spolupracovali v období 2012 – 2014 Barbora Drobíková a Pavla Švástová (metadata) a Martin Řehánek (hlavní autor softwaru CZIDLO, stejně jako většiny aplikačních příkladů použitých v metodice), Václav Rosecký a Jan Rychtář (technická část řešení). Všechny využití texty vznikly v rámci institucionálního výzkumu Národní knihovny České republiky.

DOPORUČENÁ LITERATURA

Bearman, David, Trant, Jennifer. Authenticity of digital resources : towards a statement of requirements in the research process. *D-Lib magazine* [online]. June 1998, Volume 4, Issue 6 [cit. 2014-09-01]. Dostupné z: <http://www.dlib.org/dlib/june98/06bearman.html>. ISSN 1082-9873.

Bermès, Emmanuelle. Persistent identifiers for digital resources : the experience of the national library of France. *International Preservation News* [online]. December 2006, vol. 40, s. 22–26 [cit. 2014-09-01]. Dostupné z: <http://archive.ifla.org/VI/4/news/ipnn40.pdf>. ISSN 0890-4960.

Caplan, Priscilla. The preservation of digital materials. *Library Technology Reports*. Feb/March 2008, vol. 44, issue 2, s. 4-38.

Cubr, Ladislav. Budování důvěryhodného systému trvalé identifikace digitálních dokumentů. *Knihovna*. 2010, roč. 21, č. 1, s. 23-31. Dostupné z: <http://knihovna.nkp.cz/knihovna101/10123.htm>. ISSN 1801-3252.

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

Cubr, Ladislav; Vašek, Zdeněk. Identifikátory digitálních dokumentů se zaměřením na systém URN:NBN v ČR. *Čtenář - měsíčník pro knihovny*. 2013, roč. 65, č. 6, s. 214-217. Dostupné z: <http://ctenar.svkk1.cz/clanky/2013-roc-65/6-2013/identifikatory-digitalnich-dokumentu-se-zamerenim-na-system-urn-nbn-v-cr-113-1484.htm>. ISSN 0011-2321.

Giaretta, David. *Advanced digital preservation*. New York: Springer, 2011, 510 s. ISBN 978-364-2168-086.

Gladney, Henry M. *Preserving digital information*. Berlin: Springer, 2007. 314 s. ISBN 978-3-54037-886-0.

Hakala, Juha. Document description and access - new challenges. In: *Sborník ze semináře CASLIN 2001 - Popis a zpřístupnění dokumentů : nová výzva* [online]. Praha: Národní knihovna České republiky, 2001 [cit. 2014-09-01]. Dostupné z:

<http://klement.nkp.cz/caslin/caslin01/sbornik/hakala.html>.

Hakala, Juha. *Using national bibliography numbers as Uniform Resource Names* [online]. 2001 [cit. 2014-09-01]. Internetový standard RFC 3188. Dostupné z:

<http://www.ietf.org/rfc/rfc3188.txt>.

Hakala, Juha. Persistent identifiers - an overview . In *PersID - building a persistent identifier infrastructure* [online]. 2010 [cit. 2014-09-01]. Dostupné z:

<http://www.persid.org/downloads/PI-intro-2010-09-22.pdf>.

Hilse, Hans-Werner; Kothe, Jochen. *Implementing persistent identifiers : overview of concepts, guidelines and recommendations* [online]. London: Consortium of European Research Libraries [cit. 2014-09-01]. Dostupné z: <http://nbn-resolving.de/urn:nbn:de:gbv:7-ISBN-90-6984-508-3-8>. ISBN 90-6984-508-3.

International DOI Foundation. *DOI handbook* [online]. International DOI Foundation, Updated March 17, 2014 [cit. 2014-09-01]. Dostupné z: <http://www.doi.org/hb.html>.

International ISBN Agency. *ISBN Users' manual : international edition*. Sixth Edition. London: International ISBN Agency, 2012. 31 s. ISBN 978-92-95055-02-5.

ISO 16363:2012. *Space data and information transfer systems -- Audit and certification of trustworthy digital repositories*. 1st ed. Geneve : ISO, 2012. 70 s.

ISO 14721:2012. *Space data and information transfer systems -- Open archival information system -- Reference model*. 2nd ed. Geneve : ISO, 2012. 126 s.

Kahn, Robert; Wilensky, Robert. A framework for distributed digital object services. *International Journal on Digital Libraries*. 2006, Vol. 6, Issue 2, s. 115-123.

Národní knihovna České republiky. *Příručka uživatele systému ISBN*. Praha: Národní knihovna České republiky, 2006. Revize březen 2011 [cit. 2014-09-01]. Dostupné z: <http://www.nkp.cz/soubory/ostatni/prirucka-isbi-13.pdf>.

Nestor working group long-term preservation standards. *Catalogue of criteria for assessing the trustworthiness of PI systems* [online]. Göttingen: Niedersächsische Staats und Universitätsbibliothek Göttingen, 2009 [cit. 2014-09-01]. Dostupné z: http://files.d-nb.de/nestor/materialien/nestor_mat_13_en.pdf. urn:nbn:de:0008-20080710227.

Nicholas, Nick; Ward, Nigel; Blinco, Kerry. A policy checklist for enabling persistence of identifiers. *D-Lib Magazine* [online]. January/February 2009, Volume 15, Number 1/2 [cit. 2014-09-01]. Dostupné z: <http://www.dlib.org/dlib/january09/nicholas/01nicholas.html>. ISSN 1082-9873.

Nicholas, Nick, Ward, Nigel, Blinco, Kerry. Abstract modelling of digital identifiers. *Ariadne* [online]. January 2010, issue 62 [cit. 2014-09-01]. Dostupné z: <http://www.ariadne.ac.uk/issue62/nicholas-et-al>. ISSN 1361-3200.

Palmer, William, Schlarb, Sven, Hahn, Matthias et al. Quality assured image file format migration in large digital object repositories British Library. *IPres* [online]. 2013, s. 1-4 [cit. 2014-08-29]. Dostupné z: http://purl.pt/24107/1/iPres2013_PDF/Quality%20assured%20image%20file%20format%20migration%20in%20large%20digital%20object%20repositories.pdf

Řehánek, Martin. *Identifikátory URN:NBN v prostředí českých knihoven a systém pro jejich správu*. Brno, 2012. Dostupné z: http://is.muni.cz/th/143357/fi_m/?lang=en. Diplomová práce. Fakulta informatiky, Masarykova univerzita v Brně.

Tonkin, Emma. Persistent identifiers : considering the options. *Ariadne* [online]. July 2008, issue 56 [cit. 2014-09-01]. Dostupné z: <http://www.ariadne.ac.uk/issue56/tonkin>. ISSN 1361-3200.

Vitiello, Giuseppe. Identifiers and identification systems : an informational look at policies and roles from a library perspective. *D-Lib magazine* [online]. January 2004, Volume 10,
Národní knihovna České republiky, Klementinum 190,110 00 Praha 1
www.nkp.cz; www.ndk.cz; <https://resolver.nkp.cz>

Number 1 [cit. 2014-09-01]. Dostupné z:

<http://www.dlib.org/dlib/january04/vitiello/01vitiello.html>. ISSN 1082-9873.

Jak bylo výše v textu zmíněno, další texty související s praktickým provozem systému jsou k dispozici na projektové stránce <https://code.google.com/p/czidlo/>. Jsou zde popsána i některá teoretická východiska celé architektury a další operace nepopsané v metodice, protože nejsou běžně předpokládány.