



národní
úložiště
šedé
literatury

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

Vašek, Zdeněk; Řehánek, Martin; Cubr, Ladislav,
2019

Dostupný z <http://www.nusl.cz/ntk/nusl-408011>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Licence Creative Commons Uveďte autora-Neužívejte dílo komerčně-Nezasahujte do díla 3.0 Česko

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 27.04.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz .



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

Verze 2.0 (2018)

Zdeněk Vašek – Martin Řehánek – Ladislav Cubr

*Český systém pro IDentifikaci a LOkalizaci digitálních dokumentů českého kulturního
dědictví*

*Realizováno v rámci institucionálního výzkumu Národní knihovny České republiky
financovaného Ministerstvem kultury ČR v rámci Dlouhodobého koncepčního rozvoje
výzkumné organizace*

Perzistentní identifikátor dokumentu: urn:nbn:cz:nk-004hvy

Obsah

SEZNAM OBRÁZKŮ A TABULEK	5
OPONENTI.....	5
1. ÚČEL METODIKY	6
2. URČENÍ.....	8
3. SOUVISEJÍCÍ NORMY A PŘEDPISY	8
4. TERMINOLOGIE.....	11
I. TEORETICKÁ ČÁST	15
5. IDENTIFIKÁTORY DOKUMENTŮ	15
5.1 Přehled stavu – identifikátory v praxi knihoven	15
5.1.1 Identifikátory tištěných dokumentů	15
5.1.2 Identifikátory a digitální dokumenty.....	16
5.2 Identifikátor URN:NBN	20
5.3 Východiska systému ČIDLO.....	21
II. PRAKTICKÁ ČÁST	23
6. SYSTÉM ČIDLO	23
6.1 Funkce systému.....	23
6.2 Resolver.....	24
7. PODMÍNKY ÚČASTI V SYSTÉMU ČIDLO	25
8. MODEL ROLÍ A ODPOVĚDNOSTÍ V SYSTÉMU ČIDLO.....	26
8.1 Návod pro zápis URN:NBN identifikátorů do elektronického katalogu	28
9. PŘEDMĚT IDENTIFIKACE.....	30
9.1 Model dokumentu v systému ČIDLO	30
9.1.1 Intelektuální entita.....	30
9.1.2 Digitální reprezentace/ dokument	31
9.1.3 Digitální instance.....	32
9.1.4 Vzájemné vztahy	33
9.1.5 Datový model	33
9.2 Typy identifikovaných dokumentů a možnosti rozšíření	34
III. IMPLEMENTAČNÍ ČÁST	35

10.	ZAPOJENÍ DO SYSTÉMU ČIDLO A JEHO PRAVIDLA	35
10.1	Obecné povinnosti	36
10.2	Povinnosti před přidělením URN:NBN identifikátoru (registrace registrátora)	37
10.3	Uživatelské účty	38
10.3.1	Možnosti operací při správě jmenného prostoru	39
10.4	Postup ve fázi přidělení URN:NBN identifikátoru	40
10.4.1	Příklady operací při registraci digitálního dokumentu	41
10.5	Povinnosti z hlediska zpřístupňování dokumentu	43
10.5.1	Příklad operace vložení digitální instance	44
10.6	Chybové stavy	45
11.	TYPY DOKUMENTŮ	45
11.1	Pokyny pro registrace jednotlivých typů dokumentů	46
11.1.1	MONOGRAPH	46
11.1.2	MONOGRAPH_VOLUME	52
11.1.3	PERIODICAL_ISSUE	54
11.1.4	PERIODICAL_VOLUME	57
11.1.5	ANALYTICAL	58
11.1.6	DIGITAL DOCUMENT	59
11.1.7	PŘÍKLADY SPRÁVNĚ VYTVOŘENÝCH ZÁZNAMŮ	60
11.2	Technická metadata zasílaná systému ČIDLO	61
12.	RESOLVOVACÍ FUNKCE SYSTÉMU ČIDLO	63
12.1	Citování	64
13.	HROMADNÉ SKLÍZENÍ POMOCÍ OAI-PMH	66
13.1	OAI-PMH Provider CZIDLO	67
13.2	OAI Adapter	68
14.	PRAVIDLA PRO ŽIVOTNÍ CYKLUS DOKUMENTU A IDENTIFIKÁTORU	70
14.1	Návod na postupy při slučování a rozpojování dokumentů	72
15.	KURÁTORSKÉ POSTUPY	73

15.1 OAI Adapter	74
15.2 Export URN:NBN	74
15.3 Kontrola dostupnosti DI	75
15.4 Indexace dokumentů	75
O AUTORECH.....	77
DOPORUČENÁ LITERATURA	77
SEZNAM PUBLIKACÍ, KTERÉ PŘEDCHÁZELY METODICE	80

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: Koncepční model objektů v systému ČIDLO (9.1.5)

Obrázek 5: Záznam identifikátoru v digitální knihovně (12.1)

Obrázek 6: Záznam dokumentu ve webovém rozhraní ČIDLO (12.1)

Obrázek 7: Schéma funkcí OAI Adapteru (13.2)

Tabulka 1: Role v systému ČIDLO (kapitola 8)

Tabulka 2: Doporučení na získání hodnoty „číslo ČNB“ z metadat (11.1.1)

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

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

Tabulka 5: Doporučení na získání hodnoty „druh dokumentu“ z metadat (11.1.1)

Tabulka 6: Doporučení na získání hodnoty „hlavní název“ z metadat (11.1.1)

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

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

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

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

Tabulka 11: Doporučení na získání hodnoty „místo vydání“ z metadat (11.1.1)

Tabulka 12: Doporučení na získání hodnoty „název nakladatele“ z metadat (11.1.1)

Tabulka 13: Doporučení na získání hodnoty „rok vydání“ z metadat (11.1.1)

Tabulka 14: Vzorové mapování metadat MODS na pole databáze ČIDLO (monografie) (11.1.1)

Tabulka 15: Vzorové mapování metadat MODS na pole databáze ČIDLO (svazek monografie) (11.1.2)

Tabulka 16: Doporučení na získání hodnoty „druh dokumentu“ z metadat (číslo periodika) (11.1.3)

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

Tabulka 18: Vzorové mapování metadat MODS na pole databáze ČIDLO (číslo periodika) (11.1.3)

Tabulka 19: Vzorové mapování metadat MODS na pole databáze ČIDLO (článek) (11.1.5)

Tabulka 20: Vzorové mapování metadat na pole databáze ČIDLO (digitální dokument) (11.1.6)

Tabulka 21: Doporučené mapování z metadat ve standardu MIX (11.2)

OPONENTI

1) Mgr. Zdeněk Hruška, Moravská zemská knihovna, Odbor digitalizace

2) Ing. Martin Lhoták, Knihovna Akademie věd ČR, v. v. i., Odbor informačních technologií a digitalizace

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. Cílem metodiky je poskytnout odborným pracovníkům registrátorů a jejich dodavatelských firem vysvětlení, jak služba ČIDLO technicky a především organizačně funguje a jakých pravidel je třeba se držet. 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.

Tato metodika v první verzi vznikla na základě zkušeností z provozu systému ČIDLO v letech 2012-2014, do kterého byla zapojena Národní knihovna ČR a další producenti z řad krajských a dalších knihoven a specializovaných společností zabývajících se digitalizací kulturního dědictví. Po certifikaci metodiky byla metodika těmito knihovnami i společnostmi využívána podle svého určení jako hlavní metodický dokument při užívání celoplošně určeného² unikátního perzistentního identifikátoru URN:NBN.

Zkušenosti s realizací pravidel této metodiky byly vyhodnoceny v roce 2017 a bylo rozhodnuto, že některé části metodiky je třeba doplnit a zpřesnit, aby byly uživatelsky jasnější a tudíž využitelnější. Rozšířil se počet zapojených knihoven na více než 100 a bylo přiděleno

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 Dle standardů Národní digitální knihovny.

více než 1,5 milionu identifikátorů. Byl také proveden obsáhlý výzkum využívání identifikátorů v dalších systémech v rámci zemí EU a jeho poznatky pomohly dále rozvinout metodiku pro tyto identifikátory v ČR (Cubr a kol., 2016).

Vývoj systému ČIDLO se nezastavil a vznikly nové verze softwarového nástroje. Jednotlivé nové verze softwaru nejsou samozřejmě důvodem pro vydání nové verze metodiky, nicméně rozvoj, který od roku 2014 proběhl, znamená v některých ohledech zcela nový přístup ke koncepčním bodům technického řešení, a proto by měla nová verze metodiky zohlednit i tento rozvoj. Ten přinesl mimo jiné zcela nový přístup k otázce sklizení dokumentů z digitálních knihoven a zajištění perzistentnosti URL odkazů. Další změny se týkají např. datového modelu pro OAI-PMH modul, nově jsou uživatelům k dispozici nové typy operací. Nová verze by měla reagovat i na doporučení, která vzešla ze samotného certifikačního procesu v roce 2015 a vzhledem k charakteru metodiky již nemohla být promítnuta do jejího textu. Cílem této verze je metodiku zestručnit a zpřehlednit, aby reagovala na praktické využití. Některé technické popisy a schémata byly odsunuty do externích dokumentů a z metodiky jsou pouze odkázány.

Dobu takřka čtyř let lze zároveň považovat za dostatečnou pro získání patřičného množství poznatků s praktickým využíváním metodiky. Při zpracování bylo přihlédnuto i k mezinárodnímu prostředí, zejména zkušenostem s provozem největší služby pro správu URN:NBN identifikátorů, kterou zajišťuje Deutsche National Bibliothek (DNB) a která pokrývá prostor Německa, Rakouska a Švýcarska. Každá z těchto zemí uplatňuje částečně odlišnou politiku, ale využívají společný nástroj. Zejména švýcarská pravidla pro uživatele jsou koncipována se stejným účelem jako tato metodika a byla jedním z inspiračních zdrojů. Jejich původní verze byla vydána v roce 2010 a aktualizována v letech 2014 a 2016 (e-Helvetica, 2016). Přihlédnuto bylo i k politice DNB (Ackermann, 2012).

Významná změna se týká okruhu dokumentů, pro které je tato metodika určena. Původní verze metodiky počítala s aktivním přidělováním pouze pro digitalizované dokumenty, i když koncepčně počítala i s dalšími typy. V roce 2018 existují reálné potřeby přidělování i pro digital born dokumenty a také např. pro zvukové dokumenty. I na toto rozšíření aktualizace metodiky reaguje a popisuje postupy pro jiné než digitalizované dokumenty.

Rozsáhlá změna funkcionalit a potřeba reagovat na zkušenosti z praxe tak vedly k potřebě vytvoření nové verze metodiky, která bude nástupcem původní metodiky certifikované 6. 5. 2015. Původní verze metodiky je i nadále dostupná pomocí svého perzistentního identifikátoru na URL odkazu: <https://resolver.nkp.cz/urn:nbn:cz:nk-0027gj>. I

po aktualizaci si metodika ponechá obecnou platnost, nebude tedy vázána na konkrétní verzi užívaného softwaru, nicméně některé funkce budou dostupné jen v novějších verzích a metodika na možnosti a způsob využívání konkrétních postupů poukazuje.

2. URČENÍ

Metodika je určena všem uživatelům služby ČIDLO, kteří jsou v roli registrátora nebo archivátora, případně jsou touto rolí pověřeni, dále také správčům digitálních repozitářů a digitálních knihoven paměťových institucí. 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 knihovni sbírkou) a externí producenti dat zajišťující digitalizaci. Metodika přináší postupy i pro uživatele z řad čtenářů a vědeckých pracovníků, kterým má napomoci v přístupu k obsahu digitálních dokumentů (k zjištění aktuálního umístění identifikovaných dokumentů v internetové síti) a k jeho důvěryhodnému citování. Těmto uživatelům však přináší jen základní informace a není to její primární účel. Tím je informace pro odborné pracovníky registrátorů a další odborné uživatele o způsobu fungování služby. Má jim poskytnout vhled do operací registrace, které obvykle probíhají v jejich institucích pomocí automatizovaných systémů. V případě výskytu chyby je častou příčinou chybné nastavení systému. Pro uživatele je tedy důležité mít znalost operací, které ve službě probíhají, aby byli schopni počátečního nastavení systému a následné analýzy objevené chyby. Pro pracovníky na pozicích správců digitálních repozitářů je pak nutná znalost celého životního cyklu identifikátorů. Bez popsaného workflow nemohou prohlásit identifikátor za důvěryhodný. Znalost technické specifikace je podmínkou bezpečnosti systémů spravujících kulturní dědictví. Metodika je tedy v rámci jednotlivých registrátorů určena širokému okruhu pracovníků.

Základní vymezení uživatelské komunity je následující:

- krajské a specializované knihovny,
- knihovny vědeckých ústavů, muzeí apod.
- vysokoškolské knihovny,
- producenti digitalizovaných dat kulturní povahy.

3. SOUVISEJÍCÍ NORMY A PŘEDPISY

Metodika nenahrazuje komplexní doporučení pro nakládání 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. Na tomto místě jsou proto připomenuty nejdůležitější normy a předpisy, jejichž znalost je pro správné pochopení metodiky potřebná.

Standard URN:NBN

Tento standard je základním předpisem pro syntax a užívání identifikátoru URN:NBN³. Jeho pravidla jsou implementována do systému ČIDLO, dodržováním jeho pravidel je garantováno i plnění standardu pro URN:NBN.

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, avšak nelze tento postup dlouhodobě doporučit.

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 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 - 3.7. Systém ČIDLO je připraven na změny verzí standardu, přebírá jen ta metadata, která patří k jádru popisu. Ani vydání nových verzí by nemělo mapování na

³ Dostupný z <https://www.ietf.org/rfc/rfc3188.txt> [citováno 2014-09-23].

databázi resolveru nijak omezit. Využití standardu MODS je povinné, jiné než doporučené verze standardu je předem nutno konzultovat s kurátorem systému ČIDLO.

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 doporučeno. Standard je dostupný z <http://www.loc.gov/standards/mix/> [citováno 2014-09-23].

Standardy Národní digitální knihovny

Metodika předpokládá, že digitální dokumenty identifikované pomocí URN:NBN identifikátoru budou vytvářeny nebo upravovány pomocí standardů Národní digitální knihovny (dále NDK), 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 doporučeno, není povinné, pokud to nevyžadují jiné předpisy.

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.

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.

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 vžitě 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 uchování 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 (ať již digitalizovaný či e-born) dokument je komplexní pojem, který má v rámci metodiky a informační vědy dva významy. Primárně bude pojem používán ve smyslu obecného označení digitálního objektu neboli 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 případně struktury digital born dokumentu. 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. Pojem bude používán 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 provedení dle konceptu FRBR.

Digitální instance (DI)

Umístění uživatelské kopie dokumentu v digitální knihovně, ze kterého může uživatel získat přístup k celému 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

System, ve kterém jsou uživatelské kopie dokumentů zpřístupňovány uživatelům.

Digitální repozitář

(synonymum: digitální úložiště)

System pro dlouhodobé uchovávání archivních kopií digitálních dokumentů. Tento system musí splňovat požadavky archivu OAIS.

Digitální reprezentace

Digitální reprezentace je množina počítačových souborů, 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

Konceptuální model Functional Requirements for Bibliographic Records – Funkční požadavky na bibliografické záznamy. <http://archive.ifla.org/VII/s13/frbr/frbr-cs.pdf>

Identifikátor

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

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) jeho 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.

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.

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

Viz kapitola 3.

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.

I. TEORETICKÁ ČÁST

5. IDENTIFIKÁTORY DOKUMENTŮ

5.1 Přehled stavu – identifikátory v praxi knihoven

5.1.1 Identifikátory tištěných dokumentů

Identifikátory 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 perzistentní identifikátor, tj. takový, který je trvale užíván k identifikaci právě té publikace, které byl původně přidělen (tj. nelze ho přidělit znovu). Globální trvalé identifikátory mají standardizovanou syntax a jsou přidělovány na základě pravidel, mezi která obvykle patří definice toho, čemu může být identifikátor přidělen. Například ISBN může být přidělen pouze knize na úrovni vydání, nikoliv exempláři⁴, tj. ISBN identifikuje dokument z hlediska jeho intelektuálního obsahu (a sekundárně též z hlediska specifických znaků vydání). K nezbytným předpokladům patří centrální správa těchto identifikátorů, která zajistí perzistenci a unikátnost.

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. 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

⁴ Všechny exempláře téhož vydání mají totéž ISBN.

centralizovanými hierarchizovanými systémy (mezinárodní agentura a síť podřízených 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. Každá jiná forma téhož vydání dostane jiný identifikátor ISBN. Pravidla užití mj. vymezují, že identifikátor ISBN musí být vytištěn v knize samotné, je tedy trvale spojen s touto knihou.

Tento model byl navržen tak, aby vyhovoval potřebám těch, kteří budou identifikátory ISBN využívat. Identifikuje tedy sériový produkt stejného obsahu. Naproti tomu jednoznačnou identifikaci jednotlivého exempláře vydání knihy ISBN neumožňuje.

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

Identifikace digitálních dokumentů získává na stále vyšší důležitosti, protože právě identifikátory jsou vstupní informací k objektu pro uživatele a zároveň nezbytnou součástí dlouhodobého uchovávání (Ras 2016). Perzistence identifikátoru se stává stále podstatnější součástí celého procesu ochrany digitálních dokumentů. Vývoj směřuje k „chytrým“ identifikátorům, které jsou schopny kromě identifikace poskytnout i další informace, čímž snižují potřebu dalších aplikací a databází. K takovým identifikátorům patří i URN:NBN a jeho aplikace ve službě ČIDLO.

5.1.2.1 Identifikátory

V případě digitálních dokumentů jsou možnosti využití identifikátorů oproti analogovým dokumentům daleko širší⁵. Digitální dokumenty jsou často online, tj. přístupné prostřednictvím internetové sítě. Zároveň je nutné je pojímat unikátně, každý představuje originál. Pro online dokumenty též platí, že na základě identifikátoru digitálního dokumentu je možné vyhledat nejen bibliografické údaje popisující tento dokument, ale také k němu

⁵ 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 rovině digitální reprezentace. V rámci dokumentace systému ČIDLO je digitální reprezentace označována jako digitální dokument.

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. 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 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.⁶

Mezi další rozšířené identifikační systémy patří Handle, DOI (Digital Object Identifier), PURL (Persistent Uniform Resource Locator), ARK (Archival Resource Key).⁷

Široké mezinárodní užití má systém Handle (www.handle.net), který vyvinula a provozuje

⁶ Systém využívá internetových standardů RFC 2141, 2483 a 3406. Dostupné z: <http://www.ietf.org/rfc/rfc2141.txt>; <http://www.ietf.org/rfc/rfc2483.txt>; <http://www.ietf.org/rfc/rfc3406.txt>. Sám je pak popsán ve standardu RFC 3188 - <http://www.ietf.org/rfc/rfc3188.txt>.

⁷ Populární, nicméně velmi erudovanou formou, přibližuje rozdíly webová aplikace vytvořená The Netherlands Coalition for Digital Preservation, dostupná z <http://www.ncdd.nl/en/pid-wijzer/> (cit. 2018-08-26). Aplikace nabízí rozhodnutí, který typ identifikátoru je nejvhodnější pro konkrétního uživatele z řad institucí.

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. Pro získání identifikátoru je třeba odevzdávat poměrně značné množství metadat, řídit se komplexními pravidly 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.

Mezi užívané patří i identifikátor ARK, jenž je opět postaven na bázi URL. ARK umožňuje identifikovat různé typy nejen digitálních objektů. Jeho schéma je poměrně komplikované, vedle povinné obsahuje i volitelnou část umožňující směrování.

5.1.2.2 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. 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. 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 uchování současně zachovat obsah dokumentu vč. zachování podstatných vlastností daného digitálního dokumentu.

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 je jedním z prvků datového modelu forma vydání. 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). Pro knihovny odpovědné za dlouhodobé uchování je z hlediska zachování dostupnosti nezbytné využívat takový identifikátor, který nepodléhá popsaným omezením. V mnoha ohledech platí, že omezení nejsou dána technickými nástroji, ale především organizačními opatřeními.

5.1.2.3 Digitalizace a trvalé identifikátory

Největší objem dat, u kterých je třeba zajistit dlouhodobé uchování, tvoří ve sbírkách českých knihoven digitalizáty. 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, nicméně jejich obsah zůstává stejný – je jím obraz tištěné předlohy. Perzistentní identifikátor URN:NBN musí být při generování uživatelské kopie z archivního dokumentu přebrán do metadat uživatelského dokumentu.

5.1.2.4 Trvalé identifikátory digital born dokumentů

Organizačně méně komplikovaná je identifikace trvale uchovávaných digitálních dokumentů, které již v této formě vznikly, tzv. digital born dokumentů. I u nich však platí

stejný princip, tedy deriváty archivního souboru musí dědit i jeho identifikátory, aby byl zajištěn přístup a důvěryhodnost odkazování z hlediska uživatelů.

5.2 Identifikátor URN:NBN

Pro systém ČIDLO byl zvolen identifikátor URN:NBN. V této kapitole bude představen z teoretického hlediska. 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 (RFC 3188). 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 Adresáři knihoven (dále ADR) platí, že B = sigla knihovny (šestimístný alfanumerický znakový řetězec evidovaný v ADR https://aleph.nkp.cz/F/?func=file&file_name=find-b&local_base=adr), 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 služby, 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 lidskými uživatelem. (Číslici nula „0“ a písmeno „O“ lze v některých znakových sadách snadno zaměnit). Řetězec, který vrací software, je tvořen vždy malými písmeny.

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

5.3 Východiska systému ČIDLO

Systém ČIDLO je postaven na následujících předpokladech:

1) Knihovny mají za cíl uchování digitálního dokumentu, jeho poskytnutí uživatelům a zachování pro budoucí uživatele.

2) **V případě digitalizace je obraz fyzické předlohy obsahem digitalizovaného dokumentu, který není závislý na konkrétním grafickém formátu.** Volba formátů pro archivní a uživatelské dokumenty (a do budoucna pro nástupnické archivní dokumenty a nové verze uživatelských dokumentů) je motivována primárně technickými omezeními.

3) **Předmětem identifikace v systému ČIDLO z těchto důvodů proto není obsah a forma (formát) digitalizovaného či digitálního dokumentu, ale pouze obsah předlohy. 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). V případě digital born dokumentů jsou samozřejmě exempláře shodné z podstaty.

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.

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). Ty uživatele nezajímají, protože neidentifikují obsah. Nicméně jiné identifikátory lze ve spojení s identifikátorem URN:NBN využít ke koordinaci správy dokumentů s tímtož 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, což může mít opět vliv na kvalitu obrazu. Z tohoto důvodu je vlastník dokumentu jednou z vlastností, kterou identifikátor zaznamenává. Z řečeného dále vyplývá, že dokument se stejným obsahem i vlastníkem může být identifikován opakovaně, pokud jde o jiný digitální dokument (odlišné technické nebo procesní nastavení). Stejně tak dokument se stejným identifikátorem může být využívána více vlastníky, pokud se na tom shodnou. Tento postup lze využít zejména u digital born dokumentů, kde jednou zpracovaný a identifikovaný dokument může být replikován do jiných knihoven se stejným identifikátorem.

Identifikovaný dokument musí splňovat tyto požadavky:

- a) trvalá hodnota a význam dokumentu z hlediska kulturního dědictví, v případě nejjasnosti je doporučena konzultace s kurátorem služby
- 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, případné další typy 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).

II. PRAKTICKÁ ČÁST

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 přesměrovávací 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.

Další texty související s praktickým provozem systému jsou k dispozici na projektové stránce <https://github.com/NLCR/CZIDLO>. Jsou zde popsána i některá teoretická východiska celé architektury a další operace nepopsané explicitně v metodice.

6.1 Funkce systému

Systém ČIDLO má několik funkcí, které mohou využívat uživatelé jak z řad institucí a jejich pracovníků, tak i vnější uživatelé. Funkcemi systému, které vychází z popisu využívání identifikátorů v teoretické části, jsou zejména:

- 1) Přidělování a správa unikátních perzistentních identifikátorů

Jde o primární funkci systému, identifikátory digitálních dokumentů slouží primárně pro správu digitálních dokumentů v repozitářích a prezentačních systémech. Identifikátory

jsou nezbytně nutné pro správce dat, budou využívány ke všem operacím správy a dlouhodobého uchování. Centrální služba v národním prostoru zajišťuje unikátnost a možnost koordinace při sdílení dat mezi institucemi.

2) Směrování na aktuální lokaci digitálního dokumentu

Tato služba je cílena na uživatele z řad klientů knihoven. Podstatou služby je využití identifikátoru jako části URL linku, který vždy vede na aktuální umístění digitálního dokumentu. To je dosaženo periodickým sklizením zaregistrovaných digitálních knihoven. Nástroj resolver pak k záznamu ukládá vždy platné lokace. Podrobněji jsou resolvovací funkce popsány v kapitole 12.

3) Podpora důvěryhodnosti citací

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, zůstane záznam v systému a vždy tak budou dohledatelné informace o dokumentu (identifikátorová metadata), který byl citován.

6.2 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 aktuální 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 digitální dokumenty spadající do kulturního dědictví (kartografické dokumenty, hudebniny, šedá literatura, zvukové dokumenty apod.). Registrátorem tedy může být každá instituce s patřičnými kompetencemi v oblasti správy a zpřístupňování digitálních dokumentů. V případě, že je nějaká instituce vázána povinností ukládat své dokumenty do NK ČR, nebo má záměr je zde ukládat, pak pro ni platí povinnost řídit se standardy NDK.

Od verze standardů NDK vydaných 2014 (verze 1.5 pro periodika a 1.1 pro monografie) je využívání identifikátoru URN:NBN povinné. 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 digitálního 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),
- digitální dokument není bohemikum (stejně vymezeného jako v předchozím případě)
- digitální dokument nemá trvalou hodnotu a význam z hlediska českého kulturního dědictví (v případě nejasností ohledně určení tohoto kritéria rozhodují stanoviska kurátora služby ČIDLO),
- obsah digitálního dokumentu není určen k zpřístupňování široké veřejnosti (tajné materiály, průmyslová dokumentace atd.),
- žadatel nemá záměr digitální dokument dlouhodobě uchovávat,

- žadatel nemá záměr digitální 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:**

- digitální dokumenty by měly vznikat podle standardů NDK nebo do nich být převedeny,
- v případě, že vzniknou digitální dokumenty podle jiných standardů, je jejich zapojení do služby možné, pokud bude jejich majitel schopen odevzdávat minimálně požadovaný rozsah metadat dle této metodiky,
- 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 sklizení metadat,
- zajistí uchování dokumentů, musí se zaručit, že identifikovaný dokument nesmaže.

8. MODEL ROLÍ A ODPOVĚDNOSTÍ V SYSTÉMU ČIDLO

Aktivními 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 digitálního 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, produkční systém, katalog, digitální repozitář, digitální knihovna.

Resolver je online aplikace, kterou provozuje centrální autorita z internetové adresy <https://resolver.nkp.cz>. 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.

Katalogem se rozumí elektronický, zpravidla online dostupný (OPAC) katalog registrátora, ve kterém mohou být odkazy na jeho digitální 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 digitální 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 uchování, 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 digitální dokumenty, vystupuje v systému ČIDLO v roli Archivátora.

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

Produkční 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 digitálního 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. V ojedinělých případech může jedna instituce vystupovat jako vícečetný registrátor (v případě knihoven jde např. o situace, kdy jedna instituce má více sigel). 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 (CZIDLO) 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 urgencye 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.

Pokud dojde k situaci, kdy chce či musí libovolný archivátor přijmout do svého repozitáře digitální data, která jejich majitel neopatřil perzistentním identifikátorem, je archivátor oprávněn tato data pro své interní potřeby identifikovat pomocí identifikátorů náležejících k jeho vlastnímu jmennému prostoru. S tímto identifikátorem je pak smí i zveřejňovat ve své vlastním zpřístupňovacím nástroji (digitální knihovna), pokud to bylo odsouhlaseno vlastníkem dat.

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 či RDA.

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

Pro zápis identifikátoru URN:NBN jsou určena pole 024 a 856. Primární je pole 024.⁸

⁸ Podrobněji viz metodika LICHTENBERGOVÁ, Edita et al. *Katalogizace podle RDA ve formátu MARC 21 - tištěné a elektronické monografie - katalogizace na úrovni minimálního/doporučeného záznamu* [online]. Praha: Národní knihovna ČR, 2014, aktualizace srpen 2016 [cit. 2018-09-03]. Dostupné z: <http://www.nkp.cz/o-knihovne/odborne-cinnosti/zpracovani-fondu/katalogizacni-politika/katalogizace-podle-rda-ve-formatu-marc-21-tistene-a-elektronicke-monografie-katalogizace-na-urovni-minimalniho-doporuceneho-zaznamu>.

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 digitálních monografií a jejich svazků. Pokud by měl být využit i v případě seriálů, je nutné, aby konkrétní registrátoři konzultovali s kurátorem systému, kterou úroveň chtějí odkazovat. V takovém případě lze využít i šablonu PERIODICAL (viz kapitola 11).

9. PŘEDMĚT IDENTIFIKACE

Systém ČIDLO je určen pro identifikaci statických (neměnných – viz kap. 9.1.1 – odstavec c) digitálních dokumentů bez ohledu na další typové vymezení – digital born, zvukové, digitalizované. Digitální dokument, který je předmětem identifikace, dle konceptu FRBR odpovídá úrovni provedení. Archivní i uživatelský dokument mají tentýž obsah, vlastníka a původ (vznikly v rámci téže produkce), mohou se lišit pouze konkrétním obrazovým formátem nebo např. použitým kompresním poměrem v rámci jednoho formátu. Jsou identifikovány stejným identifikátorem. V případě migrací na nástupnické verze digitálního dokumentu z důvodu zachování čitelnosti formátu existují dvě strategie, které může repozitář využít:

1) Ponechání původního identifikátoru a zachytit migraci v metadatech, aby byl schopen určit aktuální verzi.

2) Přidělit nový identifikátor pro migrovaný dokument, přičemž je doporučeno provázat nový identifikátor se starým vazbou předchůdce – následník, čímž bude zachována perzistentnost odkazů vytvořených s pomocí identifikátoru URN:NBN.

Uživatelský dokument ve formátu, ve kterém byl vytvořen při produkci, se nazývá původní uživatelský dokument. **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.

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 digitálních dokumentů je intelektuální entitou taková jednotka, kterou lze samostatně zpracovat a popsat.

Závazná pravidla:

- a) 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
- b) intelektuální entita představuje smysluplný informační celek
- vyloučeny jsou příliš malé (např. stránka knihy⁹) 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
- c) intelektuální entita představuje neměnný informační celek
- vyloučeny jsou dynamické celky (jako např. on-line zpravodajství)
 - je však možné uplatnit pro obraz dynamického celku, který je zachycen k určitému časovému úseku (např. pomocí uložení ve webovém archivu)
- d) intelektuální entita v případě digitálních dokumentů je vždy takový obsah, který je stejný na úrovni provedení (manifestation) podle modelu FRBR¹⁰
- 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
- e) intelektuální entita musí představovat informační celek, který má trvalou hodnotu a význam („*lasting value and significance*“ - definice digitálního dědictví podle organizace UNESCO¹¹)
- f) intelektuální entita má vždy samostatná bibliografická metadata.

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.

⁹ Jiný případ samozřejmě je, pokud jedna stránka reprezentuje např. jednu povídku (viz např. Borgesovy povídky)

¹⁰ http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf [citováno 2014-09-25]

¹¹ <http://unesdoc.unesco.org/images/0013/001331/133171e.pdf#page=80>. Citováno [2017-09-25]

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 obsah, 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:

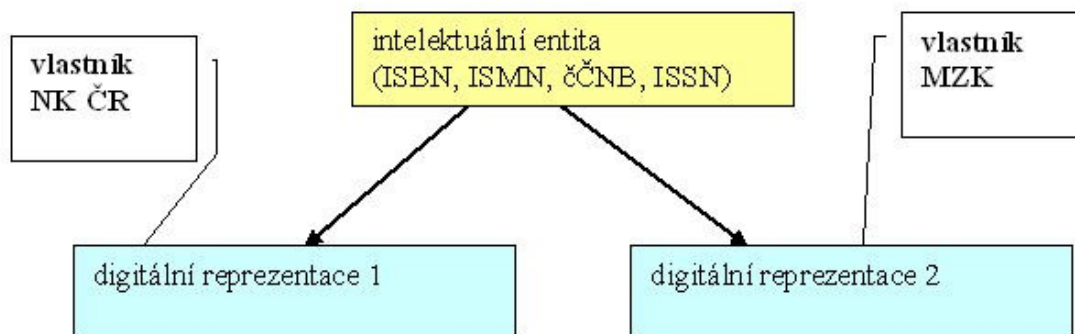
1. vstupní stránka, která se vyznačuje:
 - a) jedním URL, na němž jsou:
 - b) přístupná bibliografická, volitelně i technická metadata k uživatelskému dokumentu,
 - c) přístupné odkazy na jeden nebo více verzí uživatelského dokumentu,
2. v rámci jedné digitální knihovny existuje pouze jedna vstupní stránka.

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. Stejně tak, pokud stejný digitální dokument bude vlastnit více knihoven, mohou být označeny odlišnými identifikátory.

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.



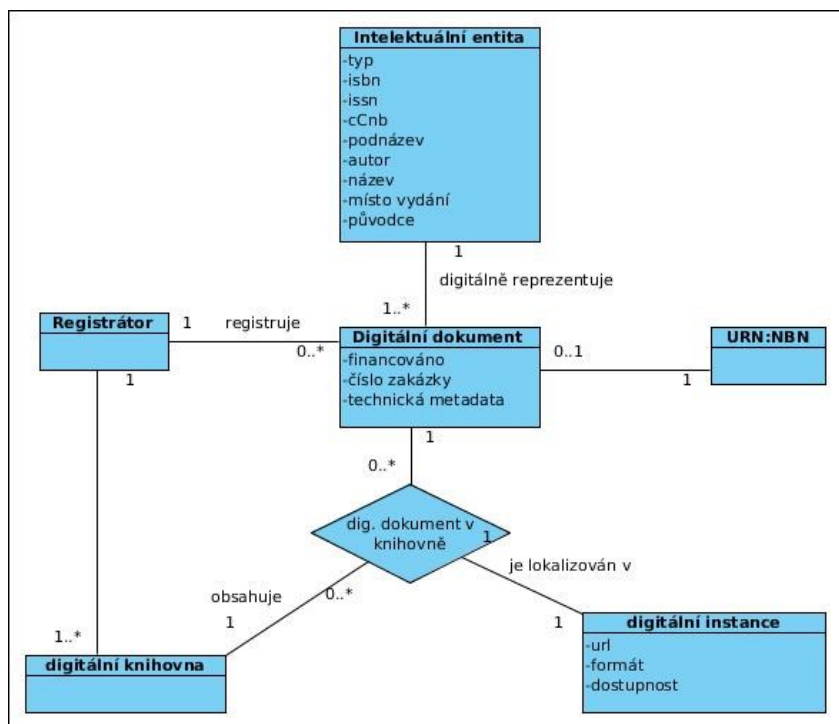
Obrázek 3: Popis vztahů mezi intelektuální entitou a digitální reprezentací.

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 vlastník, 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.

9.2 Typy identifikovaných dokumentů a možnosti rozšíření

Pomocí identifikátoru URN:NBN lze označovat libovolné digitální dokumenty, pokud splňují další parametry definované touto metodikou. Hlavní podmínkou je cíl dlouhodobého uchování a zpřístupňování dokumentu. Význam identifikace se neztrácí i při pouhém uchování, ale nelze při tom poskytnout přidanou hodnotu, kterou uživatel od služby očekává. Tato metodika má přinést návod, jak prakticky identifikátory využívat, proto se omezuje jen na ty dokumenty, pro které existují předpisy, již byly v reálném provozu vyzkoušeny – jde tedy o digitalizované i digital born dokumenty a zvukové dokumenty. Jako dokument lze označit klasické knihovní entity, jako jsou monografie a čísla periodik, ale i další dokumenty, ať již jde o články nebo studie, popsané mapy nebo např. metodiky a další. V případě zvukových dokumentů je pak identifikovanou částí to, co je při zpracování označeno jako intelektuální entita a záleží na typu zvukového záznamu a jeho nosiče, pokud záznam vznikl digitalizací. Další oblastí, ve které lze identifikaci pomocí URN:NBN rozšířit je prostředí on-line periodik, která vydávají statické dokumenty. Redakční on-line systémy lze považovat za nástroje na úrovni digitálních knihoven. Možnosti systému jsou široké, vždy je nutné konzultovat s pracovníky služby, zda je identifikátor možné využít pro danou konkrétní nestandardní entitu. Význam perzistentní identifikace v rámci uchování kulturního dědictví v elektronické podobě je tak významný, že není cílem uměle omezovat možnosti služby, pokud je rozšíření v souladu s jejím základním konceptem.

III. IMPLEMENTAČNÍ ČÁST

10. ZAPOJENÍ DO SYSTÉMU ČIDLO A JEHO PRAVIDLA

Cílem systému ČIDLO má tři hlavní funkce – viz 6.1. Jeho využití se tedy předpokládá v těchto třech oblastech. Aktuálně slouží pro distribuci perzistentních identifikátorů pro digitalizaci, která probíhá podle Standardu NDK v knihovnách České republiky. V testovacím provozu pak přiděluje identifikátory též digital born a zvukovým dokumentům. Zvláště oblast digital born nabízí do budoucna významné rozšíření – systém lze využít pro identifikaci dokumentů spadajících pod předpokládaný budoucí povinný výtisk eknih, respektive periodik. Lze ho využít též pro identifikaci vědeckých publikací, kvalifikačních prací nebo produkci on-line periodik.

Vlastníci digitálních dokumentů se musí do služby zaregistrovat, ať již sami nebo prostřednictvím pověřených zástupců (typicky společností provádějících digitalizaci službou). Registraci akceptují pravidla provozu a zavazují se je dodržovat. V případě digital born publikací je registrátorem obvykle knihovna, která dokumenty spravuje nebo je získala pro dlouhodobé uchování. Pro všechny typy dokumentů platí pravidlo, že dokumenty, které vstupují do repozitáře pro dlouhodobé uchovávání musí být opatřeny identifikátorem od vlastníka. Pokud tomu tak není, je správce repozitáře oprávněn přidělit identifikátory jako by byl sám vlastníkem těchto dokumentů.

Po zaregistrování dostanou registrátoři přidělen přístup do služby (uživatelský účet) a mohou ji využívat při zpracování svých dokumentů. Nejčastěji se tak děje automaticky, pomocí API, které poskytuje resolver služby ČIDLO. Přidělení se děje pomocí zaslání skupiny vybraných metadat, služba za ně systému registrátora obratem vrátí perzistentní identifikátor. Ten je zapsán do metadat dokumentu. Tím je realizována první oblast činnosti ČIDL a, tedy přidělování perzistentních identifikátorů a udržování jejich databáze. **Služba garantuje, že byl přidělen globálně unikátní identifikátor.** K záznamu konkrétního identifikátoru jsou připojena zasláná metadata. Ta následně slouží k ověřování citací. Tedy pokud by v budoucnu dokument byl z digitální knihovny stažen, uživatel má pomocí identifikátoru, který si zapsal, stále možnost dohledat bibliografická data digitálního dokumentu. Díky tomu může prokázat, že jím citovaný dokument existoval. Nelze samozřejmě již zobrazit vlastní obsah.

Po zveřejnění identifikovaného dokumentu je třeba do ČIDLO dodat automatizovaně či ručně URL odkaz na daný dokument. Případně periodicky spouštět sklizeň digitální

knihovny, aby aplikace sama přiřadila URL odkaz (digitální instanci) k záznamu identifikátoru.

Tyto kroky je třeba provést v době vzniku digitálního dokumentu, respektive v době jeho přijetí do systému dlouhodobého uchování a zveřejnění. V další fázi životního cyklu digitálního dokumentu je třeba dodržovat pravidla, jak s identifikovaným dokumentem nakládat. Pravidla a postupy, které je třeba dodržovat, jsou popsány v následujících kapitolách této metodiky.

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

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 (respektive metadata daného dokumentu) je možné měnit jen za vymezených podmínek popsanych 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), respektive mít tento záměr. Dokument může v rámci procesů životního cyklu i zaniknout.
- 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** (s výjimkou, kdy jde o přímou přejímku dat k dlouhodobému uchování, kterým

jejich vlastník nepřidělil identifikátor URN:NBN).

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

- I. Registrované dokumenty nemusí být zveřejněny ihned, ale v určité časové lhůtě.
- II. 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.
- III. **Správce repozitáře potřebuje identifikovat předané a dosud neidentifikované dokumenty opatřit identifikátory.** V takovém případě by měl oslovit předávající, aby identifikaci zajistil. Pokud to není možné, smí se souhlasem kurátora systému ČIDLO užít k identifikaci vlastní jmenný prostor, aby se předešlo rizikům při dlouhodobém uchovávání.

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 urnmbn@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 dalšími 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 obdrží přístupové údaje k účtu a následně 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://github.com/NLCR/CZIDLO>. Především nastaví správné rozpoznávání druhů digitálních dokumentů systému ČIDLO. Registrátorovi lze udělit práva k ručnímu zadávání požadovaných metadat pomocí webového formuláře (v tomto případě nejde o technické, ale administrativní opatření).
- 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. **Heslo k účtu může být uživatelem změněno.**

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 zajišťuje kontrolu

duplicitního přidělení identifikátoru sám resolver v okamžiku zpracování registrace dokumentu. V případě využití rezervace aplikace kontroluje, zda nebyl identifikátor použit dvakrát a případně zamítne dokončení registrace.

10.3.1 Možnosti operací při správě jmenného prostoru

Uživatel (registrator či jeho pověřený zástupce) může v rámci svého jmenného prostoru provádět různé operace, které mu slouží pro potřeby úspěšného vykonávání správy nad přiděleným prostorem. Software resolveru je postupně vyvíjen a inovován. Souběžně s tímto vývojem vznikají nové verze aplikačního rozhraní. Pro zaručení zpětné kompatibility je paralelně udržováno několik starších verzí API. **Tato metodika doporučuje pravidelně aktualizovat nastavení softwaru registratorů tak, aby odpovídal nejnovější dostupné verzi API.** Pro správnou funkčnost celého systému doporučuje metodika v době svého vydání využívat rozhraní APIv3 a vyšší. V případě využívání APIv2 není možné garantovat plnou funkčnost systému pro všechny popsané příklady. Aktuálně doporučené je rozhraní APIv5. Od APIv4 jsou dostupné vybrané odpovědi nejen ve formátu XML, ale i v JSON.

Správci systémů registratorů by měli nejméně jednou ročně kontaktovat kurátora resolveru a zkontrolovat nastavení svého systému. Příklady použité v této metodice budou různorodé, aby dostatečně ilustrovaly možnosti, které lze použít. Budou však dodržovat doporučení, aby šlo nejméně o příklady podle APIv3. Některé z operací jsou dostupné i přes webové rozhraní pro manuální správu.

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: <https://resolver.nkp.cz/api/v4/registrars/nk?format=xml> nebo <https://resolver.nkp.cz/api/v4/registrars/nk?format=json>). 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/v4/registrars/kód_registrátora/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 možných operací, které API resolveru nabízí, je dostupný na projektové stránce <https://github.com/NLCR/CZIDLO/wiki/API>.

Podrobné popisy operací pro jednotlivé verze API jsou k dispozici zde:

APIv3

<https://docs.google.com/spreadsheets/ccc?key=0Ag5aMq4LaXVcdFZ3MjBOb2s2UIZtb0ZqbV93WWJ4ZGc>

APIv4

<https://docs.google.com/spreadsheets/d/1QT1dLjsjZrXzqdv-TqD18UTrCmzpQ24-ilv7-X1RbvY>

APIv5

<https://docs.google.com/spreadsheets/d/1JBWhkdek0AOT-uo3QC7GApkI6yP6GrvMLJkAbVW1LM>

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.

Procesy dostupné i přes webové rozhraní (v závislosti na uživatelské roli):

- a) Úprava informací o registrátorech a archivátorech a jejich digitálních knihovnách
- b) Zobrazení statistik o přiřazování identifikátorů a o resolvování
- c) Vytvoření nového záznamu
- d) Správa účtů včetně uživatelské změny hesla
- e) Správa procesů (viz kapitola 15)

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 registraci zasláním požadovaných metadat. Je doporučeno, aby registrovači 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.

Pokud jde o zaslání metadat, systém ČIDLO pevně požaduje takový rozsah metadat, který umožní jednoznačnou identifikaci dokumentu. Povinný rozsah je blíže specifikován v kapitole 11. Nicméně je doporučeno zaslání co nejširšího popisu, aby byla podpořena

funkce zvýšení důvěryhodnosti citací digitálních dokumentů.

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

```
<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) 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. **Metadatový popis zaslaný resolveru již registrátor dále měnit nesmí (v případě nutnosti je třeba kontaktovat kurátora a domluvit s ním postup).** V případě, že je tento postup nutný, přikročí se k deaktivaci dokumentu a k vytvoření nového záznamu a to buď pomocí vytvoření vazby předchůdce – následník, což je preferovaný způsob, anebo deaktivací a vytvořením zcela nového dokumentu. Tento postup není doporučen.

Resolver umožňuje i operaci (od rozhraní APIv3) doplnění záznamu. Tato operace je určena k zaslání dodatečných informací a to v případě, kdy daný element nebyl v původní žádosti vůbec vyplněn. Jde o jediný případ, kdy lze změnit podobu záznamu již zaregistrovaného dokumentu. Typickým příkladem je, pokud si registrátor pro své dokumenty nechá až dodatečně přidělit identifikátor čČNB. Následně může provést doplněný metadatového balíčku o tento element a operaci doplnění záznamu v resolveru. Je to však povoleno jen pro doplnění chybějícího údaje, nesmí dojít k přepsání již existujícího.

10.4.1 Příklady operací při registraci digitálního dokumentu

URN:NBN identifikátor je přidělován digitálnímu dokumentu operací *Register 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>

nebo

<http://resolver.nkp.cz/api/v4/digDocRegistration.xsd>

nebo

<http://resolver.nkp.cz/api/v5/digDocRegistration.xsd>

Schéma pro import dokumentu je dostupné např. zde:

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

<http://resolver.nkp.cz/api/v5/digInstImport.xsd>

V následujících příkladech jsou využity ukázky konkrétních dokumentů a postupů, jak provést registraci podle obou automatizovaných typů registrace – tedy pomocí on-line registrace (By_Resolver) nebo rezervací (By_Reservation). Jednotlivá pole v databázi budou vysvětlena v kapitole 11 a to včetně doporučení jakými hodnotami tato pole lze plnit. Jde o příklady strukturovaných dat pro strojové zpracování přístupné i lidským uživatelům. Běžný uživatel z řad veřejnosti bude přistupovat přes grafické webové rozhraní, kde bude v kontaktu pouze se záznamem v podobě formuláře a nikoli zde prezentovanými příklady, které mají soužití pro lepší pochopení struktury záznamu. Kompletní přehled uživatelských operací, které lze provádět, je odkázán v kapitole 10.3.1 a liší se podle verze aplikačního rozhraní.

1. Registrace digitálního dokumentu – mód BY_RESOLVER

CZIDLO přiřadí volné (nerezervované) URN:NBN.

Request

POST <http://resolver.nkp.cz/api/v4/registrars/boa001/digitalDocuments>

Tělo requestu bude obsahovat importní xml s metadaty intelektuální entity. Podrobnější popis requestu lze nalézt na následujícím odkazu:

https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#registrace-dd--m%C3%B3d-by_resolver

Příklad platí pro request dle APIv3. Pro další verze aplikačního rozhraní lze nalézt příklady operací zde: <https://docs.google.com/spreadsheets/d/1QT1dLjsjZrXzqdv-TqD18UTrCmzpQ24-ilv7-X1RbvY>

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>

Podrobnosti: <https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#request-5>

Seznam rezervovaných URN:NBN získáte tímto dotazem:

Request

GET <http://resolver.nkp.cz/api/v4/registrars/mzk/urnNbnReservations>

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

Request

POST <http://resolver.nkp.cz/api/v4/registrars/boa001/digitalDocuments>

3. Mód BY_REGISTRAR

Vedle obou zmíněných postupů existuje i možnost přidělování registrátorem, kdy on volí podobu identifikátoru, ale tato není běžně pro registry 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>

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 přidělený URN:NBN identifikátor.

Příklad volání:

Request

POST

<http://resolver.nkp.cz/api/v4/resolver/urn:nbn:cz:mzk-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, do které 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

Každá digitální instance v systému má vlastní id, které je interní a vzniká v rámci systému.

Digitální instance lze vkládat a deaktivovat jak automatizovaně pomocí API, tak manuálně pomocí webového rozhraní. Pomocí tohoto dotazu lze získat informaci o všech digitálních instancích daného dokumentu

10.6 Chybové stavy

Při komunikaci s resolver může docházet k zaslání chybných údajů. Na jednotlivé typy chyb reaguje resolver chybovými hlášeními a jejich kódy. Tyto kódy se liší podle verze aplikačního rozhraní. Pro správnou interpretaci chybového hlášení je proto třeba použít podrobné popisy reakcí aplikačního rozhraní, které je odkazováno v kapitole 10.3.1.

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 touto metodikou **zakázáno používat**, lze to změnit na základě odůvodněné žádosti registrátora. Typ dokumentu je určující pro volbu šablony, podle které následně probíhá registrace. Typ dokumentu již nelze v budoucnu změnit a to ani při operaci vytváření následníka. Základní schéma resolveru předpokládá využití pro klasické dokumenty ve fondech knihoven, což vyhovuje i pro digital born monografie a periodika. Typ Thesis je určen pro kvalifikační práce vysokých škol a typ Analytical slouží pro záznam článků v periodicích, podmínkou je přiřazení k nadřazenému periodiku. Typ Other slouží pro ostatní možné dokumenty včetně zvukových.

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 mohou objevit hodnoty jako např. “mapa”, “grafika”, „zvukový dokument“.**

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 dle standardu NDK může obsahovat několik URN:NBN

identifikátorů (vždy však platí, že v jedné úrovni metadatového popisu smí být jen jeden platný URN:NBN identifikátor a libovolné množství neplatných). 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, respektive u jiných typů to, co je považováno za intelektuální entitu. URN:NBN identifikátor získaný pro tuto intelektuální entitu musí být zapsán do metadatového popisu na úroveň předpokládanou Standardy NDK pro konkrétní typ dokumentu. 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).

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

Typ 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, pro digital born 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 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 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

Mapování metadat pro request na aplikační rozhraní ČIDLO:

Následující části popisují mapování pro získání hodnot k plnění polí v databázi systému ČIDLO.

Tabulka 2: 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">	

Tabulka 3: 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">	

Jiný identifikátor jako např. vnitřní identifikátory lokálníchází mohou být převedeny do tohoto pole.

Tabulka 4: 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	
DC kvalifikovaný	<code><dc:identifier type=""></code>	
DC nekvalifikovaný	<code><dc:identifier></code>	
MODS	<code><mods:identifier type=""></code>	

Tabulka 5: 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	
DC kvalifikovaný	<code><dc:type></code>	
DC nekvalifikovaný	<code><dc:type></code>	
MODS	<code><mods:typeOfResource></code> ; <code><mods:genre></code>	

Tabulka 6: 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
MARC21	245 \$a	
DC kvalifikovaný	<code><dc:title></code>	
DC nekvalifikovaný	<code><dc:title></code>	V poli dc title se spojuje title a subTitle ve formátu „title: subTitle“

MODS	<mods:titleInfo><mods:title>	
------	------------------------------	--

Tabulka 7: Doporučení na získání hodnoty „autor“ z metadat

Název pole	Název podpole	Název pole v databázi
Autor	Autor	primaryOriginator/Autor
Název pole	Šablona metadat	Povinnost
Autor	<pre><xs:element ref="r:primaryOriginator" minOccurs="0" maxOccurs="1" /> <xs:simpleType name="originatorTypeType"> <xs:enumeration value="AUTHOR" /></pre>	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	100 \$a, 100 \$b, 100 \$c, 100 \$d	
DC kvalifikovaný	<dc:creator>	
DC nekvalifikovaný	<dc:creator>	
MODS	<mods:name><mods:namePart>	

Tabulka 8: 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	<pre><xs:simpleType name="originatorTypeType"> <xs:enumeration value="CORPORATION" /></pre>	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	110 \$a, 110 \$b	
DC kvalifikovaný	<dc:creator>	
DC nekvalifikovaný	<dc:creator>	
MODS	<mods:name><mods:namePart>	

Tabulka 9: 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	<code><xs:simpleType name="originatorTypeType"> <xs:enumeration value="EVENT" /></code>	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	111 \$a, 111 \$n, 111 \$c	
DC kvalifikovaný	---	
DC nekvalifikovaný	---	
MODS	<code><mods:name><mods:namePart></code>	

Tabulka 10: 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	<code><xs:element name="otherOriginator" type="r:otherOriginatorType" minOccurs="0" maxOccurs="1" /></code>	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	700 \$a, 700 \$b, 700 \$c, 700 \$d	
DC kvalifikovaný	<code><dc:contributor></code>	
DC nekvalifikovaný	<code><dc:contributor></code>	
MODS	<code><mods:name><mods:namePart></code>	

Tabulka 11: 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	<code><xs:element name="place" type="r:placeType" minOccurs="0" maxOccurs="1" /></code>	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	260 \$a	
DC kvalifikovaný	<code><dc:publisher></code>	
DC nekvalifikovaný	<code><dc:publisher></code>	
MODS	<code><mods:originInfo><mods:place></code>	

Tabulka 12: 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	<xs:element name="publisher" type="r:publisherType" minOccurs="0" maxOccurs="1" />	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	260 \$b	
DC kvalifikovaný	<dc:publisher>	
DC nekvalifikovaný	<dc:publisher>	
MODS	<mods:originInfo><mods:publisher>	

Tabulka 13: 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	<xs:element name="year" type="xs:integer" minOccurs="0" maxOccurs="1" />	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	260 \$c	
DC kvalifikovaný	<dcterms:issued> ; <dcterms:created> ; <dcterms:modified>	
DC nekvalifikovaný	<dc:date>	
MODS	<mods:originInfo><mods:dateIssued>	

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

Tabulka 14: 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']"		
documentT	z metsu pomocí getType		

ype			
digitalBorn	False, true		
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 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 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 s názvem 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 mapování z metadat MODS pro tento typ, které metodiku doporučuje, vypadá takto:

Tabulka 15: 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)		
volumeTitl e	//mods:titleInfo/mods:partNumber	//mods:titleInfo/mods:partName	
ccnb	//mods:identifier[@type='ccnb']		
isbn	//mods:identifier[@type='isbn']		
otherId	//mods:identifier[@type='isrc']"		
documentT ype	z METS pomoci metody getType()		
digitalBorn	False/true (pro digitalizované dokumenty vždy false)		
primaryOri ginator	//mods:name[@type='personal' and @usage='primary']//mods:namePart[not(@type='date')]	//mods:name[@type='corporate']	//mods:name[@type='conference']
otherOri ginator	//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. Zde je třeba vyplnit číslo, které jednotlivý svazek periodika má. Pokud není číslo k dispozici, vyplní se co nejpřesnější

datum vydání tak, aby umožnilo jednoznačnou identifikaci dokumentu. Tedy, pokud jde o měsíčník, postačí název měsíce, ve kterém vychází, jde-li o deník nebo týdeník, tak ideálně vyplnit přesné denní datum. Ve sporných případech je třeba se řídit Pravidly popisu NDK.

- 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, pro digital born dokumenty 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 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, 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, v tomto případě má datum vydání jen orientační význam, nemusí být vyplněno, v případě, že vyplněno je, obsahu informaci o roku vydání.

Popis polí, která jsou odlišná od ostatních typů dokumentů:

Tabulka 16: 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	<code><xs:element name="documentType" type="r:documentTypeType" minOccurs="0" maxOccurs="1" /></code>	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>	

Tabulka 17: 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	<xs:element name="issn" type="r:issnType" minOccurs="0" maxOccurs="1" />	Ne
Mapování		
Standard	Pole (Tag)	Poznámky
MARC21	022 \$a	
DC kvalifikovaný	<dc:identifier type="urn">	
DC nekvalifikovaný	<dc:identifier>	
MODS	<mods:identifier type="issn">	

Doporučené mapování:

Tabulka 18: 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 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/standardy-digitalizace/metadata> [citováno 2018-09-25].

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
- 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, pro digital born dokumenty vždy 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, 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, 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, nemusí odpovídat názvu ročníku

Popis je obdobný jako u předchozího typu.

11.1.5 ANALYTICAL

Typ se používá pouze pro analytika, tedy pro samostatné články.

- 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, 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, 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 19: 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 20: Vzorové mapování metadat na pole databáze ČIDLO (digitální dokument)

Název pole	hodnota
archvierId	Viz popis
urnnbn	///mods:identifiek[@type='urnnbn']
registrarScopeIdentifiers	vložen následující registrarScopeIdentifier
registrarScopeIdentifier_type	uuid
registrarScopeIdentifier_value	Hodnota cdmId
financed	Zdroj financování, slovní popis
contractNumber	Číslo dodávky/ smlouvy

Hodnoty registrarScopeIdentifier_type a _value jsou ponechány 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.

11. 1.7 PŘÍKLADY SPRÁVNĚ VYTVOŘENÝCH ZÁZNAMŮ

1) PERIODICAL_ISSUE, XML

<https://resolver.nkp.cz/api/v4/resolver/urn:nbn:cz:nk-003alt?format=xml>

2) PERIODICAL_ISSUE, JSON

<https://resolver.nkp.cz/api/v4/resolver/urn:nbn:cz:nk-003alt?format=json>

3) MONOGRAPH, XML

<https://resolver.nkp.cz/api/v4/resolver/urn:nbn:cz:nk-003aet?format=xml>

4) MONOGRAPH, JSON

<https://resolver.nkp.cz/api/v4/resolver/urn:nbn:cz:nk-003aet?format=json>

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 21: 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/CompressionScheme
	ratio	//mix/BasicDigitalObjectInformation/Compression[0]/CompressionRatio
color	model	//mix/BasicImageInformation/BasicImageCharacteristics/PhotometricInterpretation/colorSpace
	depth	//mix/ImageAssessmentMetadata/ImageColorEncoding/BitsPerSample/bitsPerSampleValue
iccProfile		//mix/BasicImageInformation/BasicImageCharacteristics/PhotometricInterpretation/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 neboli 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=[show,redirect,decide,NEUVEDENO]** a **format=[html, xml, NEUVEDENO]**, které využívají verze resolveru založené na APIv3 a starších. U novějších verzí je k dispozici již jen jediný parametr **format=[xml, json, NEUVEDENO]**.

Bez upřesňujících parametrů se resolver 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.

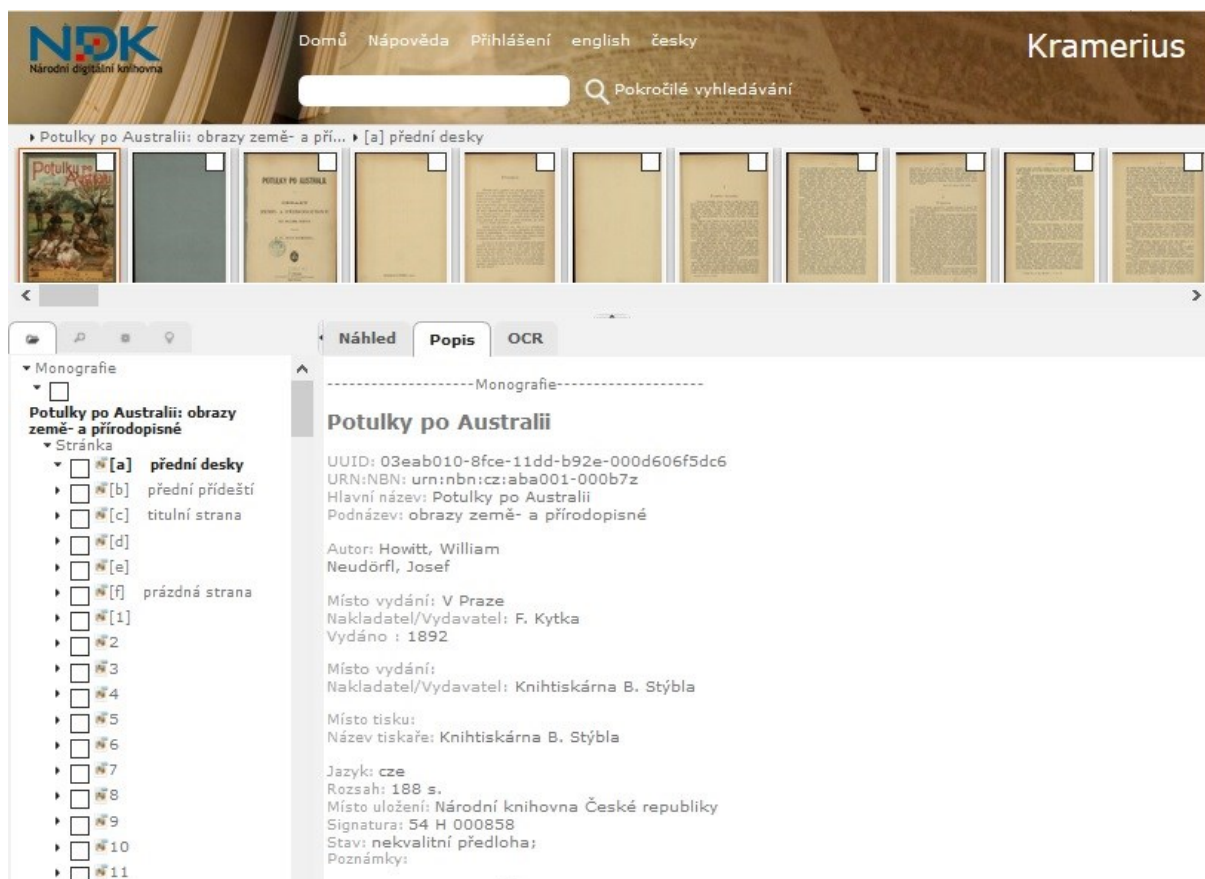
Uživatel znající hodnotu identifikátoru může získat přístup k žádanému dokumentu třemi způsoby, pokud bude využívat identifikátor a možnosti, které nabízí systém ČIDLO a spolupracující systémy.

- 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:aba001-000b7z>. 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 využije odkaz v katalogu (pole 856), bude přesměrován primárně na digitální instanci náležející do knihovny provozované vlastníkem katalogu. Např. když budou v katalogu odkazy typu <https://resolver.nkp.cz/urn:nbn:cz:aba001-000b7z> 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.

Resolver je schopný též tzv. reverzního resolvování, tedy převedení URL adresy na URN:NBN záznam.

12.1 Citování

Jedním z cílů činnosti služby ČIDLO je podpora důvěryhodnosti citací digitálních dokumentů. Podrobněji byla funkce popsána v kapitola 6.1. Citování pomocí identifikátoru URN:NBN může mít dvě podoby. Lze citovat za použití samotného identifikátoru. V tom případě uživatel obvykle použije klasický URL odkaz, ale přidá k němu informaci o identifikátoru digitálního dokumentu. Ten by měl být viditelný ve webovém rozhraní příslušné digitální knihovny. Např. v prostředí digitální knihovny Kramerius provozované Národní knihovnou ČR je identifikátor zapsán v záložce Popis.



Obrázek č. 5: Záznam identifikátoru v digitální knihovně

V případě, že se uživatel chce vyhnout citování URL dokumentu zcela, může použít již zmíněnou verzi s pomocí webu resolveru. Dokument cituje s pomocí URL odkazu resolveru, který je složen z adresy resolveru a identifikátoru dokumentu <https://resolver.nkp.cz/urn:nbn:cz:aba001-000b7z>. Tento odkaz zajišťuje resolvování přímo do digitální knihovny a případně i do všech dalších zaregistrovaných digitálních knihoven. Zvyšuje se tím šance, že se uživatel dostane k požadovanému obsahu, i kdyby byl na původní adrese nedostupný.

Strukturu záznamu v resolveru ukazuje následující obrázek. Rozsah záznamu je závislý na informacích, který do služby vložil registrátor, který odpovídá za jejich rozsah a správnost. Služba jako povinný vyžaduje jen minimální rozsah popisných metadat. I v případě deaktivace dokumentu zůstává celý záznam uložený a vyhledatelný. Uživatel se vždy dostane k základní popisným informacím.

Informace Pravidla Vyhledávání Registrátoři Statistiky

urn:nbn:cz:nk-001zzz Hledat

▣ Krajiny vnitřní a vnější (monografie)

[zobrazit v ČNB](#)

název: Krajiny vnitřní a vnější
 podnázev: texty o paměti krajiny; smysluplném bobrovi, areálu jablkového štrúdlu a také o tom, proč lezeme na rozhlednu
 ČČNB: onb001571412
 ISBN: 8073830427
 digital born: ne
 autor: Cílek, Václav
 jiný původce: Šejn, Miloš
 místo vydání: Praha
 rok vydání: 2005
 vloženo: 23.10.2015 10:33:12

▣ digitální dokument [XML](#) [JSON](#)

▣ urn:nbn:cz:nk-001zzz [XML](#) [JSON](#)

registátor: Národní knihovna České republiky
 jiný archivátor: Národní knihovna České republiky [podrobnosti](#)
 vloženo: 23.10.2015 10:33:13

▣ technická metadata

▣ registrar-scope identifikátory

uid: ebb57800-5c30-11e5-bf4b-005056827e51

<http://www.digitaln knihovna.cz/mzk/uid/uid:ebb57800-5c30-11e5-bf4b-005056827e51> [XML](#) [JSON](#)

<http://kramerius4.nkp.cz/search/handle/uid:ebb57800-5c30-11e5-bf4b-005056827e51> [XML](#) [JSON](#)

▣ deaktivovaná instance [XML](#) [JSON](#)

url: <http://krameriusndktest.mzk.cz/search/handle/uid:ebb57800-5c30-11e5-bf4b-005056827e51>
 formát: jpg;jp2;pdf
 dostupnost: chráněno autorskými právy
 digitální knihovna: Kramerius 4 [podrobnosti](#)
 vloženo: 23.10.2015 15:53:53
 deaktivováno: 11.11.2017 05:38:54

Obrázek č. 6: Záznam dokumentu ve webovém rozhraní ČIDLO

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*. Úplná specifikace OAI-PMH je dostupná na <https://www.openarchives.org/OAI/openarchivesprotocol.html>.

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*.

Sety a selektivní sklizně

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.

Kontrola toku

Data provider zpravidla obsahuje tisíce a milióny záznamu, proto není praktické, aby operace ListRecords a ListIdentifiers vrátili opravdu všechny záznamy, resp. hlavičky. Namísto toho vrací maximálně několik záznamů a tzv. resumption token. To je řetězec, který lze použít pro další dotaz ListRecords/ListIdentifiers, tentokrát už bez dalších parametrů, kromě samotného resumptionToken. Tento dotaz vrátí další balík záznamů a další resumption token. Tak se pokračuje, dokud není sklizeno vše, co bylo prvním dotazem případně omezeno (*set*, časový rozsah).

13.1 OAI-PMH Provider CZIDLO

CZIDLO realizuje OAI-PMH Data provider pomocí aplikace oaiPmhProvider. Funkci 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/web>. Více lze nalézt na projektové stránce systému ČIDLO <https://github.com/NLCR/CZIDLO/wiki/OAI-PMH-data-provider>.

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

Dostupné metadatové formáty jsou pro všechny záznamy stejné, konkrétně vlastní formát CZIDLO a povinný formát oai_dc (verze Dublin Core).

13.2 OAI Adapter

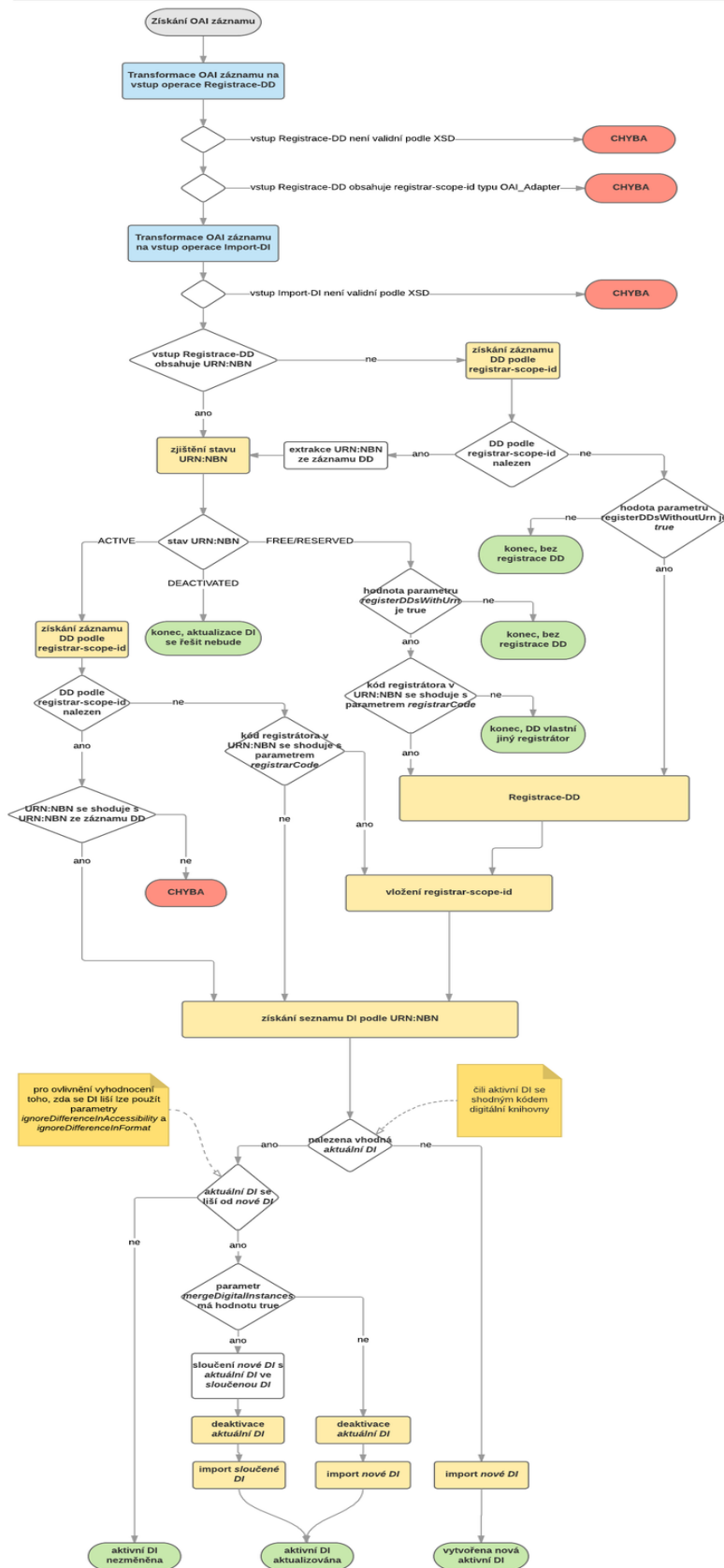
Dále systém ČIDLO disponuje nástrojem OAI Adapter. OAI Adapter je nástroj umožňující registrátorům vkládat a udržovat aktuální odkazy na zveřejněné instance digitálních dokumentů s přiřazeným URN:NBN identifikátorem. Je určen pro ty digitální knihovny, které disponují OAI-PMH rozhraním. Skládá se z těchto 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 změnilo. Více informací o OAI Adapteru je k dispozici na projektové stránce <https://github.com/NLCR/CZIDLO/wiki/OAI-Adapter> a v uživatelském manuálu <https://github.com/NLCR/CZIDLO/raw/master/documents/OaiAdapter%20-%20u%C5%BEivatelsk%C3%BD%20manu%C3%A1l.pdf>. OAI Provider lze spouštět z webového rozhraní v sekci Procesy (viz kapitola 15).



Obrázek č. 7: Schéma funkcí OAI Adapteru

14. PRAVIDLA PRO ŽIVOTNÍ CYKLUS DOKUMENTU A IDENTIFIKÁTORU

Jeden digitální dokument, jemuž byl 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 dokument vznikl dle pravidel digitalizace NDK). Ani jeden z nich pak již nesmí být měněn. Popisná metadata uživatelské verze musí odpovídat popisným metadatům archivního balíčku v poměru 1:1.

V případě budoucích formátových migrací uskutečněných v digitálním repozitáři s digitálním dokumentem¹² 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 ČSN ISO 16363 a ČSN ISO 14721 a dalších osvědčených postupů užívaných v odborné komunitě. Technická 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 signifikantní 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 a registrovat dokument s novým identifikátorem URN:NBN bez vytváření následnické vazby. Pokud registrátor dokument trvale smazal nebo ho trvale zneprístupnil ve všech instancích, pak musí deaktivovat jeho záznam 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).**

Vymazání uuid 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.

¹² Například migrace souborů v JPEG 2000 do nějakého budoucího nového grafického formátu.

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>
```

Odstranění uuid a dalších interních identifikátorů lze provést i pomocí webového rozhraní.

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

Requesty jsou popsány podrobně zde:

<https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#vlo%C5%BEen%C3%ADaktualizace-identifik%C3%A1toru>

Pokud digitální dokument identifikátor tohoto typu zatím neměl, dojde k vložení nového.

Pokud identifikátor existoval, je jeho hodnota změněna (aktualizace).

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 metadatach 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

Podrobnosti v podobě struktury zasláního příkazu do aplikačního rozhraní jsou popsány zde:

<https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#p%C5%99edch%C5%AFdci-an%C3%AIsledn%C3%ADc%C3%AD>

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 resolování 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)

Podrobnosti v podobě struktury zasláního příkazu do aplikačního rozhraní jsou popsány zde:

<https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#nahrazen%C3%AD-digit%C3%A1ln%C3%ADho-dokumentu-nap%C5%99-zm%C4%9Bna-v-metadatech>

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

Podrobnosti v podobě struktury zasláního příkazu do aplikačního rozhraní jsou popsány zde:

<https://github.com/NLCR/CZIDLO/wiki/API-v3---p%C5%99%C3%ADklady#rozd%C4%9Blen%C3%AD-digit%C3%A1ln%C3%ADho-dokumentu-ve-dva>

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.

15. KURÁTORSKÉ POSTUPY

Kurátor systému ČIDLO je klíčovou osobou v rámci celého systému. Nicméně část kurátorských operací pro svůj jmenný prostor mohou vykonávat i samotní registrátoři, respektive jejich pověřeni pracovníci. Ti mohou vystupovat v rolích kurátorů svých jmenných prostorů, nemají tedy právo zřizovat další účty, vytvářet účty institucí atd. Mohou nicméně

využívat nástroje pro tvorbu přehledů o zaregistrovaných dokumentech, které lze dále pro účely správy dat třídit a zpracovávat. Tyto postupy umožňují kvalitativně lepší využívání identifikátorů URN:NBN pro správu digitálních dokumentů v repozitářích. Kromě nástrojů repozitářů může každý správce vlastního prostoru využít nástroje resolveru.

Přihlášený uživatel systému může přes webové rozhraní v záložce „PROCESY“ nastavit a spustit několik procesů.

15.1 OAI Adapter

Spouští procesy AOI Adapteru a umožňuje jeho nastavení. Toto webové rozhraní dovoluje nahrát XSLT šablony, které budou využity pro operace „Transformace pro operaci Registrace digitálního dokumentu“ a „Transformace pro operaci Import digitální instance“.

15.2 Export URN:NBN

Pomocí příkazu „EXPORT URN:NBN“ lze spustit procesy exportů URN:NBN záznamů podle uživatelem zvolených kritérií. Administrátor může po stisknutí tlačítka „plánovat proces“ zaškrtnout jeden nebo více parametrů z daných kategorií, na základě kterých je exportována (do formátu csv) výsledná množina záznamů z databáze CZIDLO, která splňuje všechny zadané parametry. Některé kategorie (jejich parametry) nemusí být vybrány. Lze využít následující kategorie:

- **Časové období** (datum registrace) – Tato kategorie umožňuje vybrat časové období, v rámci kterého byly URN:NBN záznamy registrovány. Parametry zahrnují tyto položky: rok, měsíc, den. Tyto položky se uplatňují ve dvou podkategoriích – počáteční okamžik časového období, konečný okamžik časového období. V rámci této kategorie musí být logicky vždy zadány všechny parametry.
- **Registrátor** – Parametry v této kategorii odpovídají jednotlivým registrátorům, označených kódem registrátora. Administrátor může označit jednoho nebo více registrátorů, ke kterým se URN:NBN záznamy vztahují, a to v okně se seznamem kódů registrátorů řazeným podle abecedy. V rámci této kategorie musí být vždy vybrán alespoň jeden parametr (tj. alespoň jeden registrátor).
- **Druh dokumentu (typ IE)** – Parametry této kategorie jsou povolené typy dokumentů (klasifikované podle typu intelektuální entity, v primární databázi atribut entityType), kterým lze přidělit identifikátor URN:NBN: *monografie, svazek monografie, ročník periodika, číslo periodika, vysokoškolská závěrečná práce, analytikum, jiná*

intelektuální entita. V rámci této kategorie musí být vždy vybrán alespoň jeden parametr (tj. alespoň jeden typ dokumentu).

- **Záznamy postrádající klíčové identifikátory** – Parametry této kategorie jsou tři klasické identifikátory – Číslo České národní bibliografie (čČNB), ISBN, ISSN. Vybrání parametru znamená požadavek takové množiny výpisů URN:NBN záznamů, která neobsahuje dané identifikátory. Absence těchto identifikátorů může někdy poukazovat na významnou chybu v produkci digitalizovaných dokumentů. Parametry této kategorie mají povahu konjunkce.
- **Počet digitálních instancí** – Tato kategorie obsahuje jediný parametr, jehož označením bude vygenerován výpis obsahující u každého URN:NBN záznamu počet digitálních instancí k danému záznamu (tj. k danému digitálnímu dokumentu označenému identifikátorem URN:NBN), tedy na kolika internetových adresách URL je daný dokument zpřístupňován uživateli.
- **Stav URN:NBN** – Tato kategorie obsahuje tři parametry, jež se vzájemně vylučují: všechny, aktivní, neaktivní. Aktivní URN:NBN je takové, jež označuje funkční dokument (dokument, který je buď uložen v digitálním repozitáři, nebo současně též v digitální knihovně, a který není ve vztahu k jinému dokumentu předchůdcem). Neaktivní URN:NBN je takové, které označuje nefunkční dokument (takový, který je ve vztahu k jinému dokumentu předchůdcem, nebo takový, který zanikl).

15.3 Kontrola dostupnosti DI

Nově software CZIDLO umožňuje proces kontrol dostupnosti DI. Jde o operaci spouštěnou z webového rozhraní. Cílem tohoto procesu je kontrola, zda jsou všechny zaregistrované URL odkazy stále platné. Procesy spouštěné pomocí OAI Adapteru nejsou totiž schopny rozeznat již neaktuální URL odkazy a výsledek sklizně může být proto chybný. Právě funkcionality „kontrola dostupnosti“ umožňuje ověření platnosti všech odkazů, nejen těch, které nalezne Adapter.

15.4 Indexace dokumentů

Tento proces slouží k indexaci pro Solr index. Ten se používá pro vyhledávání přes webové rozhraní. Ten proces slouží k doplnění/aktualizaci indexu v nestandardních situacích, jako je přechod ze starší verze CZIDLO (indexace všeho), zpětného řešení výpadků (proběhly importy dat, ale indexace z nějakého důvodu ne). Indexace pro každý importovaný záznam spouští běžně, nicméně asynchroně, aby indexace neovlivnila výsledek importu a nebrzdila

samotný import. Uživatel má možnost nastavit časové období, pro které chce indexaci spustit. V případě problému při importu se doporučuje krátký časový interval bezprostředně předcházející zjištěnému problému.

O AUTORECH

1) Metodika verze 1.0 (2014)

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.

2) Metodika verze 2.0 (2018)

Autory aktualizace metodiky jsou Mgr. Martin Řehánek a PhDr. Zdeněk Vašek, Ph.D. za spolupráce s PhDr. Mgr. Ladislavem Cubrem, Ph.D. Všechny využití texty vznikly v rámci institucionálního výzkumu Národní knihovny České republiky.

DOPORUČENÁ LITERATURA

Ackermann, Uta, Berner, Christiane, Elbert, Natalie et al. *Policy for issuing URNs in the urn:nbn:de namespace* [online]. Version 1.0. Leipzig; Frankfurt am Main: Deutsche Nationalbibliothek, 2012 [cit. 2018-08-28]. Dostupné z: <http://nbn-resolving.org/urn:nbn:de:101-2013121001>

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.

Bellini, Emanuele, Bergamin, Giovanni, Messina, Maurizio et al. NBN:IT The Italian trusted persistent identifier infrastructure. *International Journal of Knowledge and Learning* [online]. 2014, vol. 9, no. 4, p. 1-17 [cit. 2015-07-20]. ISSN 1741-1017. Dostupné z: <http://dx.doi.org/10.1504/IJKL.2014.069536>

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.

e-Helvetica: URN-Handbuch: Zusammengetragen und erstellt durch die URN-Koordinationsstelle der Schweizerischen Nationalbibliothek. Update 2016 [cit. 2018-09-10]. Dostupné z: <https://www.nb.admin.ch/snl/de/home/nb-professionell/e-helvetica/infos-fuer-anbieter-innen/urn.html>

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.

LICHTENBERGOVÁ, Edita et al. *Katalogizace podle RDA ve formátu MARC 21 - tištěné a elektronické monografie - katalogizace na úrovni minimálního/doporučeného záznamu* [online]. Praha: Národní knihovna ČR, 2014, aktualizace srpen 2016 [cit. 2018-09-13].

Dostupné z: <http://www.nkp.cz/o-knihovne/odborne-cinnosti/zpracovani-fondu/katalogizacni-politika/katalogizace-podle-rda-ve-formatu-marc-21-tistene-a-elektronicke-monografie-katalogizace-na-urovni-minimalniho-doporuceneho-zaznamu>

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, Wiliam, 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

Ras, Marcel. Building a Future for our Digital Memory A National Approach to Digital Preservation in The Netherlands. *iPRES 2016. Proceedings*. Swiss National Library. [cit. 2018-09-01], Dostupné z: http://www.ipres2016.ch/ipr16.organizers-congress.org/frontend/indexa37c.html?folder_id=353

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, Number 1 [cit. 2014-09-01]. Dostupné z: <http://www.dlib.org/dlib/january04/vitiello/01vitiello.html>. ISSN 1082-9873.

Wannenwetsch Oliwer, Tim A. Majchrzak, On constructing persistent identifiers with persistent resolution targets, *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, Gdansk, 2016, [cit. 2018-09-01], Dostupné z: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7733372&isnumber=7733200>

SEZNAM PUBLIKACÍ, KTERÉ PŘEDCHÁZELY METODICE

Cubr, Ladislav; Zdeněk Vašek. 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.

Cubr, Ladislav; Zdeněk Vašek, Martin Řehánek. *Analýza VŠ repozitářů z hlediska zapojení do systému ČIDLO*. Národní knihovna České republiky 2014. Interní studie.

Cubr, Ladislav, Zdeněk Vašek et al. Srovnání vybraných národních identifikačních systémů užívajících identifikátory URN: NBN. *ProInflow: časopis pro informační vědy* [online], 2016, 8(1) [cit. 2018-09-05]. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/1220/1490>.

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.

Řehánek, Martin. *Softwarové nástroje pro podporu URN:NBN ve vybraných evropských zemích*. Národní knihovna České republiky 2015. Interní studie.

Vašek, Zdeněk, Martin Řehánek. *Analýza možností rozšíření působnosti systému ČIDLO o on-line periodika*. Národní knihovna České republiky 2016. Interní studie.

Vašek, Zdeněk, Cubr, Ladislav a kol. *Metodika pro přidělování a správu životního cyklu unikátních perzistentních identifikátorů digitálních dokumentů podle standardu URN:NBN* [online]. Praha, 2014, 87 s. [cit. 2016-04-12]. URN:NBN urn:nbn:cz:nk-0027gj. Dostupné z: <https://resolver.nkp.cz/urn:nbn:cz:nk-0027gj>

Vašek, Zdeněk. Чешская реализация URN: NBN системы. *Infolib*, 2017, 1. ISSN 2181-8207. Dostupné z: <http://natlib.uz/uzb/NewArrivals/ReadMore/6> [cit. 2018-09-01]