

Cubr, Ladislav; Lodrová, Iveta; Řehánek, Martin; Vašek, Zdeněk

Srovnání vybraných národních identifikačních systémů užívajících identifikátory URN:NBN

ProInflow. 2016, vol. 8, iss. 1, pp. 13-53

ISSN 1804-2406

Stable URL (handle): <https://hdl.handle.net/11222.digilib/135195>

Access Date: 17. 02. 2024

Version: 20220831

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

Srovnání vybraných národních identifikačních systémů užívajících identifikátory URN:NBN*

***The Comparison of the selected European national identification systems using
identifiers URN: NBN***

Ladislav Cubr, Iveta Lodrová, Martin Řehánek, Zdeněk Vašek

Národní knihovna České republiky

Recenzenti:

PhDr. Hana Vochozková

Ing. Alojz Androvič, CSc.

Abstrakt

Studie se zabývá identifikačními systémy zemí EU založenými na standardu URN:NBN. Pro analýzu byly zvoleny systémy ve Finsku, Itálii, Nizozemsku a v Německu (německý systém poskytuje služby i pro Rakousko a Švýcarsko). Analýze jsou podrobena jednotlivá technická řešení a národní odlišnosti jednotlivých modelů implementace standardu URN:NBN. Studie v úvodu přináší přehled identifikačních systémů pro digitální dokumenty a podrobně popisuje standard URN:NBN. V hlavní části studie pak jsou na základě zvolených kategorií (objekt identifikace, účastníci systému, model syntaxe a pravidla používání) porovnávány jednotlivé národní identifikační systémy s implementací identifikačního systému ČIDLO (systém pro URN:NBN v ČR). Cílem studie je poukázat na možnosti a omezení identifikačního systému využívajícího URN:NBN a především na jeho roli v procesech dlouhodobého uchování a zpřístupňování digitálních dokumentů z oblasti národního kulturního dědictví.

Klíčová slova: *RFC 3188, URN:NBN, resolver, perzistentní identifikátory, digitální dokumenty, systémy identifikace, Česká republika, Finsko, Itálie, Německo, Nizozemí*

Abstract

The study deals with the identification systems based on EU standards URN:NBN. For the analysis were selected systems in Finland, Italy, the Netherlands and Germany (The German system provides services for Austria and Switzerland too). The paper analyzes the technical solutions and national diversity of individual cases implementations of the URN: NBN standard. The study in the introduction gives an overview of the

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

identification system for digital documents and describes in detail the standard URN:NBN. The main part of the study is based on a comparison of selected categories (object of identification, participants in the system, the syntax model and the rules of system) of various national identification system, including systems ČIDLO (Czech system based on standard URN:NBN). The study aims to present the possibilities and limitations of the system using standard URN:NBN and to emphasize its role in the processes of long-term preservation and accessibility of digital documents of national cultural heritage.

Keywords: RFC 3188, URN:NBN, resolver, Persistent Identifiers, Digital Documents, Identification Systems, Czech Republic, Finland, Germany, Italy; Netherlands

Úvod

Jednoznačná a především trvalá identifikace je v současném digitálním světě stále potřebnější. S rostoucím počtem digitálních objektů je jednoznačná identifikace nezbytným předpokladem pro další využití i správu těchto objektů. Identifikační systémy se stávají čím dál komplexnějšími a multifunkčními. Identifikátory digitálních objektů plní obvykle tři funkce: identifikace (jednoznačné označení) referentu¹, zprostředkování online zpřístupnění digitálního objektu a mezinárodní platnost. Jelikož jsou tyto funkce úzce spjaty s provozem internetové sítě, není překvapivé, že nejznámějším digitálním identifikátorem současnosti je URL adresa. Ta však neplní ani jednu z výše uvedených funkcí dlouhodobě: internetové adresy jsou známy svou nestabilitou. Jeden z nedávných výzkumů například udává, že běžná doba platnosti URL odkazu je přibližně sto dní [Schweizerische, 2014, s. 6]. Ačkoliv se v případě stránek paměťových institucí obvykle setkáme s delší dobou platnosti internetového odkazu, pravidlo zůstává stejné: URL adresa není dobrý perzistentní (trvalý) identifikátor. Objekty digitálního kulturního a vědeckého dědictví však potřebují perzistentní identifikátory, které je trvale identifikují a pomocí nichž jsou zpřístupňovány uživatelům.

Perzistentní identifikátory vyžadují, aby existoval komplexní identifikační systém pro jejich přidělování a správu. Takové identifikační systémy existují již dlouho pro oblast tištěných dokumentů, připomeňme systémy ISBN (ISO 2108:2005) nebo ISSN (ISO 3297:2007). Základem těchto identifikačních systémů je soubor pravidel (pravidla pro syntax identifikátoru jsou pouze jedním z mnoha takovýchto pravidel) a vymezení rolí jednotlivých účastníků (registrační agentury).

Všechny úspěšné identifikační systémy pro perzistentní digitální identifikátory jsou založeny na tezi, že identifikátor je identifikátorem digitálního dokumentu, nikoliv lokace, na níž je momentálně daný dokument zpřístupňován uživatelům. Z tohoto důvodu tyto identifikační systémy vyžadují navíc speciální technickou infrastrukturu. Jejím základem je resolver, což je služba, která provádí tzv. dereferenci, což je „proces, při kterém je identifikátor vstupem (požadavkem) do síťové služby, která jako odpověď poskytne specifický výstup (zdroj, metadata atd.).“ [Bellini, 2008, s. 2] Dereferenci (v angličtině často označováno také termínem „resolving“) lze provádět s identifikátory jakýchkoliv dokumentů (digitálních i tištěných). Dereferencí identifikátoru tištěné knihy však můžeme získat maximálně bibliografický záznam s informací o tom, kde přesně je kniha uložena. Správa identifikátorů a jejich dereference jsou dvě relativně nezávislé skupiny procesů, které ale musí fungovat nad společnou databází. Podle toho pak vypadají konkrétní softwarová řešení. V některých případech vše funguje v rámci jedné aplikace, často jsou ale tyto dvě funkce oddělené. Využívaným řešením je správa a přidělování identifikátorů v oddělených databázích jednotlivých jmenných podprostorů a následná registrace v centrální databázi určené primárně pro dereferenci identifikátorů.

V praxi identifikačních systémů pro perzistentní digitální identifikátory však jde především o možnost získat dokument samotný. Resolver je v tomto případě tedy chápán jako online služba, která je umístěna na určité (ideálně neměnné a v dané komunitě uživatelů dostatečně známé) internetové adrese, funguje přes protokol HTTP a na základě zadání perzistentního identifikátoru provádí v první řadě přesměrování na URL adresu, na níž se daný digitální dokument aktuálně nalézá; případně může vykonat některé další operace (např. zobrazí metadata, která se vážou k digitálnímu dokumentu nebo identifikátoru). Součástí této technické infrastruktury musí být nutně mechanismy, které zajistí, že jednou přidělený perzistentní identifikátor nebude přidělen znovu a že resolver bude vždy schopen získat informaci o aktuální lokaci digitálního objektu. Tím je zajištěno, že identifikátor zprostředkovává online přístup k digitálnímu dokumentu, i když URL adresa se může měnit.

¹ Referentem je zde myšleno to, co je identifikátorem označováno. Nemusí jít jen o objekt, ale též o událost apod.

Zároveň platí, že perzistentní identifikátor zůstává ze své podstaty stále stejný i při změně URL adresy identifikovaného objektu.

Jak již bylo poznamenáno v mnoha studiích, perzistence identifikátorů závisí na perzistenci identifikačního systému (srv. např. [PILIN, 2007]). Identifikační systém sám musí fungovat trvale. Perzistence identifikátorů je tak spíše otázkou toho, jak dobře je identifikační systém administrativně zajištěn, což předpokládá odpovídající organizační, personální a finanční zdroje.

Dále je třeba upozornit na potřebu uchovávat v identifikačním systému také související metadata (vztahující se k identifikátoru nebo referentu). Jak upozornil Juha Hakala ve svém příspěvku pro konferenci CASLIN 2001, pro to, aby identifikační systémy „fungovaly dobře, musí být všechny systémy identifikátorů doplněny souvisejícími metadaty, např. ISSN by nebylo k ničemu bez metadat, která seriál, jemuž byl tento identifikátor přidělen, popisují. Skutečně: standard ISSN vyžaduje, aby seriál, kterému má být přiděleno ISSN, byl zkatalogizován. Systém ISBN však (bohužel) takový požadavek nemá.“ [Hakala, 2001a, s. 23]

Na tomto místě je třeba velmi poděkovat pracovníkům národních knihoven a dalších zúčastněných institucí, kteří nám poskytli cenné informace a konzultace k jednotlivým implementacím identifikačních systémů a to ať už pomocí emailové komunikace nebo při osobním setkání. Informace nám poskytli Esa-Pekka Kesitalo a Janne Pietarilla (Finsko), Emanuele Bellini (Itálie), Uta Ackermann, Nicole von der Hude a Kadir Karaca Koçer (Německo) a Rob de Bruin (Nizozemsko). Bez jejich spolupráce by tato studie nemohla vzniknout.

Hlavní identifikační systémy pro perzistentní digitální identifikátory

V polovině 90. let 20. století vytvořili Robert Kahn a Robert Wilensky nový model architektury digitální knihovny. Jeho součástí je digitální objekt, jehož složkami jsou data a klíčová metadata (key-metadata). Druhá jmenovaná složka zahrnuje globální perzistentní identifikátor, specificky identifikátor Handle. [Kahn, 2006]² V téže době byl spuštěn Systém Handle (The Handle System, <http://handle.net/>) pro správu těchto identifikátorů.³ Součástí Systému Handle je mezinárodní resolver, který dosud stále běží na své původní internetové adrese (<https://hdl.handle.net/>), což z něj činí nejstarší úspěšný identifikační systém pro perzistentní digitální identifikátory.

Na technické infrastruktuře Systému Handle byl později vystavěn nový identifikační systém.⁴ Nese název Systém DOI (ISO 26324:2012; The DOI System, <http://www.doi.org/>).⁵ Také systém DOI také provozuje svůj vlastní resolver na trvalé adrese (<https://dx.doi.org/>).

² Tato studie je citována podle její pozdější přetištěné verze, která je obsahově totožná s původní verzí. Původní verze je stále dostupná na adrese <http://www.cnri.reston.va.us/k-w.html>.

³ Systém Handle a zmíněná architektura digitální knihovny byly vyvinuty americkou výzkumnou organizací CNRI (Corporation for National Research Initiatives), a to v rámci projektu CS-TR (Computer Science Technical Reports), který byl financován vládní agenturou Ministerstva obrany Spojených států amerických DARPA (Defense Advanced Research Projects Agency).

⁴ Systém DOI vyvíjela v letech 1996-1998 taktéž organizace CNRI, a to ve spolupráci s mezinárodními vydavatelskými asociacemi. Jeho první verze byla spuštěna v roce 2000.

⁵ „DOI“ je zkratka pro „Digital Object Identifier“, ovšem ve smyslu „digitální identifikátor objektu“. Tím chtěli tvůrci systému poukázat na skutečnost, že s identifikátorem se pracuje prostřednictvím digitálních technologií. Sám identifikátor DOI sice převážně označuje právě digitální objekty, ale může být také užít k identifikaci nedigitálních entit, například obchodních transakcí nebo osob.

V roce 2001 byl pak představen standard pro další identifikační systém, URN:NBN, známý jako RFC 3188. [Hakala, 2001b]⁶

Identifikační systémy DOI a URN:NBN jsou v současnosti nejužívanějšími systémy pro perzistentní identifikátory označující objekty digitálního kulturního a vědeckého dědictví. Identifikátor DOI je ve větší míře užíván delší dobu než URN:NBN a převažujícím typem referentu, který je jím označován, jsou články odborných periodik vydávané v digitální podobě nebo zpřístupňované jako digitalizované verze tištěných periodik.

Systém DOI má hlavního (mezinárodního) správce, kterým je Mezinárodní nadace pro DOI (International DOI Foundation, IDF). Tento systém využívá technickou infrastrukturu Systému Handle tak, že nadaci IDF je přiřazen jeden z jeho jmenných prostorů (s číslem „10“). V rámci tohoto prostoru je pak (již v rámci Systému DOI) delegována pravomoc přidělovat identifikátor DOI několika registračním agenturám. Tyto agentury přidělují identifikátory daným registrátorům, zpravidla vydavatelům periodik.

Pravidla Systému DOI představují vysoce komplexní strukturu, vytvořenou s cílem podporovat komerční vydavatelskou činnost, podobně jako je tomu u systému ISBN (ovšem s větší centralizací a systémovými možnostmi). Přidělení identifikátoru DOI je mj. podmíněno dodáním požadovaných metadat popisujících identifikovaný referent (velký důraz je kladen na záznam práv duševního vlastnictví) a tím, že referent musí být zpřístupňován prostřednictvím úvodní popisné stránky (obsahující bibliografické údaje a provozované daným vydavatelem) a ze které je dostupný (zpravidla ke stažení) identifikovaný digitální článek nebo jiný typ dokumentu.⁷ Vlastní resolver Systému DOI slouží pro dereferenci identifikátorů DOI, ale dokáže pracovat i s identifikátory Handle.⁸ V případě identifikátorů DOI primárně přeměrovává na zmíněnou úvodní popisnou stránku.

Identifikační systémy pro perzistentní identifikátory, které se staly úspěšnými (Systém Handle, Systém DOI, systém URN:NBN), se jimi staly přinejmenším ze dvou důvodů. Zaprvé proto, že přidělují identifikátory digitálním dokumentům, nikoliv jejich lokacím. Zadruhé pak proto, že tyto systémy samy jsou stále udržovány a rozvíjeny svými správci. Bez trvalého spravování a vývoje identifikačních systémů by jimi přidělené digitální identifikátory pozbyly smyslu. Dále je dobré poznamenat, že není jistě náhodou, že identifikátory přidělované v těchto systémech se využívají především pro objekty digitálního dědictví (vědecké publikace, digitalizované knihy apod.), tedy objekty, u kterých se očekává, že budou trvale uchovávány jako součást národního odkazu. Trvale identifikovat objekty pomíjivé povahy by bylo mrháním vynaloženého úsilí. Nicméně je důležité upozornit, že v odborné literatuře se teprve nedávno objevilo zdůraznění toho, že perzistentním identifikátorem by měl být identifikován jen objekt, u kterého bude zaručeno jeho dlouhodobé uchovávání, a že pomíjivost identifikovaného objektu snižuje smysl celého identifikačního systému. [Bellini, 2014]

Zajímavým důsledkem výše uvedeného je, že perzistentní identifikátory musejí být zachovány, i když digitální dokument přestane být online dostupný nebo když referent zanikne.⁹

⁶ URN je zkratka pro „Uniform Resource Name“, NBN pro „National Bibliography Number“. RFC (Request for Comments) je dokument vydávaný Společností pro internet (The Internet Society) a Komisí pro technickou stránku internetu (Internet Engineering Task Force).

⁷ Dostupný většinou pouze pro oprávněné uživatele za úplaty.

⁸ Resolver Systému Handle dokáže rovněž provádět dereferenci obou těchto identifikátorů.

⁹ K zániku dokumentu digitálního dědictví by nemělo docházet, ale v praxi nevyhnutelně někdy dochází a bude docházet k případům, kdy se tak stane

Systém URN:NBN

Zavádění systému URN:NBN mělo pomalejší průběh. Ačkoliv byl navržen (jako standard)¹⁰ již na počátku milénia, masověji se rozšířil až v posledních letech. Návrh standardu pro URN:NBN byl sice iniciován ze strany Konference knihovníků evropských národních knihoven (Conference of the European National Librarians, CENL, <http://www.cenl.org/>) [Hakala, 2001b], CENL však původně zvažovala stát se jednou z registračních agentur Systému DOI. Konsorcium tří evropských národních knihoven se dokonce v letech 2002-2008 stalo členy Systému DOI v roli pozorovatele. V roce 2007 však pracovní skupina pro perzistentní identifikátory, kterou založila právě CENL, doporučila evropským národním knihovnám, aby namísto Systému DOI přidělovaly identifikátory URN:NBN a zakládaly své národní resolversy pro jejich dereferenci. [Altenhoener, 2007]¹¹

Systém URN:NBN je registrovaný v rámci internetového identifikačního systému URN.¹² Před jeho zavedením existovaly identifikátory NBN jakožto perzistentní identifikátory určené k použití výhradně národními knihovnami, a to zejména pro identifikaci těch publikací, které nemají číslo ISBN nebo podobný identifikátor, nebo pro identifikaci bibliografických metadat popisujících takové dokumenty. Nejdelší tradici má identifikátor NBN patrně ve Finsku (pod názvem „F-code“). Zdejší národní knihovna jej přidělovala katalogizačním lístkům již od počátku 20. století.¹³ V České republice byl identifikátor NBN (jako „číslo ČNB“) zaveden až v roce 2010.¹⁴ Jedním z důvodů jeho zavedení byla potřeba jednoznačně propojit tištěné dokumenty, které neměly v katalogizačním záznamu žádný trvalý identifikátor, s jejich nově vznikajícími digitalizovanými verzemi.¹⁵ Zajímavostí je, že pro české knihy byl již v 50. letech 20. století užíván československý identifikační systém, do jisté míry analogický mnohem později zavedenému systému ISBN. [Pistorius, 2011, s. 57]¹⁶ Identifikátory tohoto systému byly vytištěné v knize a na přebalu a užívány až do konce 80. let 20. století. Žel však nejsou zaneseny do katalogizačních záznamů.

Jedinečnost identifikátorů NBN je však zaručena maximálně na národní úrovni. Naproti tomu identifikátor URN:NBN je mezinárodně jedinečný (což byl jeden z důvodů, proč byl navržen) a je určen pro digitální objekty, které jsou online a tedy je lze prostřednictvím identifikátoru zpřístupnit internetovým uživatelům. Podle standardu pro URN:NBN spadá přidělování identifikátorů URN:NBN výhradně do pravomoci národní knihovny dané země. Ta však může zavést taková pravidla, která umožní, aby byl identifikátor URN:NBN užíván i jinými institucemi v zemi. K tomu také v praxi dochází již od samých počátků přidělování tohoto identifikátoru.¹⁷

Je třeba dodat, že systém URN:NBN není jednolitý mezinárodní systém, jako jsou systémy DOI nebo ISBN. De facto jde o soubor jednotlivých národních identifikačních systémů, které spojuje dodržování základních pravidel předepsaných standardem pro URN:NBN. Mezi tato

¹⁰ Přesněji řečeno, dokument RFC 3188 je de facto standard. V dokumentu se uvádí: „*It does not specify an Internet standard of any kind.*“ Nicméně pro potřeby této práce jej budeme označovat jako „standard“.

¹¹ Bylo mj. konstatováno, že samotné členství v Systému DOI je pro knihovny velmi nákladné.

¹² Cílem internetového identifikačního systému URN je zaručit univerzální a jedinečnou identifikaci zdrojů, která bude nezávislá na jejich lokaci. Systém URN je spravován korporací ICANN (Internet Corporation for Assigned Names and Numbers, <https://www.icann.org/>).

¹³ Později byl přidělován záznamům ve formátu MARC. Tento identifikátor nalezneme v katalogu finské národní knihovny v poli 015.

¹⁴ <http://www.caslin.cz/spoluprace/sluzby/ccnb/?searchterm=%C4%8Dnb>

¹⁵ Digitalizovaným verzím přiděluje Národní knihovna ČR od roku 2012 identifikátor URN:NBN. Identifikátor je přidělován během produkce digitalizátu.

¹⁶ Identifikátor se skládal ze tří částí tvořených čísly a oddělených pomlčkou. První část tvořilo dvojčíferné číslo nakladatele, druhou část pořadové číslo knihy v daném nakladatelství a třetí poslední dvojčíslí letopočtu vydání.

¹⁷ Nejznámějším příkladem byl severský projekt DIVA, na kterém se podílelo několik skandinávských univerzit a který vytvořil mechanismus pro přidělování identifikátorů URN:NBN univerzitním publikacím.

pravidla patří povinnost v zemi dané národní knihovny provozovat alespoň jeden resolver pro dereferenci identifikátorů URN:NBN a zajistit, aby tyto identifikátory byly jedinečné a v souladu s předepsaným modelem pro syntax.

Konkrétní znakový řetězec v rámci jmenného prostoru tohoto identifikátoru¹⁸ se musí skládat ze dvou částí: prefixu a řetězce NBN. Tyto části musejí být vždy odděleny spojovníkem. Prefix může mít dvě varianty. V první variantě obsahuje pouze kód země podle normy ISO 3166-1 alpha-2¹⁹ (například „fi“ pro Finsko).²⁰ Model pro celý identifikátor URN:NBN této první varianty je následující:

URN:NBN:<kód země>-<řetězec NBN>

V druhé variantě je prefix tvořen tak, že je národní prostor užitím dvojtečky dále členěn na jmenné podprostory. V tomto případě národní knihovna rozděluje tento podprostor důvěryhodným institucím, kterým přiřazuje kód jmenného podprostoru (například „jyu“ je kódem finské univerzity Jyväskylän yliopisto).²¹ Tento kód pak tedy označuje danou instituci jakožto registrátora identifikačního systému. Model pro celý identifikátor URN:NBN této druhé varianty je následující:

URN:NBN:<kód země>:<kód jmenného podprostoru>-<řetězec NBN>

V praxi existují i případy dvojstupňového členění národního prostoru (například v Nizozemsku, viz dále v textu). Druhá varianta syntaxe přidává národní knihovně povinnost udržovat registr těchto kódů a zveřejňovat je na webu. Samotný řetězec NBN lze dále členit pomocí spojovníku. Pro syntax obecně platí, že identifikátor URN:NBN musí obsahovat pouze alfanumerické znaky anglické abecedy.²² Jeho délka není omezena.

Nad rámec těchto pravidel pak mají národní knihovny volnou ruku v tom, jaká zpřesňující kritéria si pro svou variantu identifikačního systému URN:NBN zvolí. V případě identifikátoru URN:NBN je perzistence závislá na fungování národních knihoven. Jak dodává Hakala, „*perzistenci kódů NBN jakožto identifikátorů zajišťuje perzistence národních knihoven a jimi udržovaných informačních systémů,*“ [Hakala, 2001b] což je podstatný moment celého systému URN:NBN. Ve všech zemích, které ho zavedly, jde o komplexní systém, který je sice provozován na konkrétním technologickém řešení, ale základní podmínkou pro jeho funkčnost, tedy persistentní identifikaci, je dodržování pravidel v národním jmenném prostoru. Ke kontrole, ale i specifikaci pravidel, stejně jako koordinaci celé sítě je též nezbytné stabilní personální zázemí. Tvůrci systému URN:NBN proto logicky předpokládali, že z hlediska stability institucí jsou právě národní knihovny nejlepší volbou pro správu tohoto systému, protože samy o sobě představují dostatečně trvalý prvek systému a obvykle také mají dostatečnou autoritu dohlížet na fungování systému a i po personální stránce jsou ho schopny nejnázne zajistit.

Jak již bylo řečeno, pro celý systém URN:NBN platí pravidla určující model syntaxe identifikátoru a některá další základní omezení. Konkrétní užití standardu stejně jako technické zabezpečení identifikačního systému zůstává v kompetenci správců jednotlivých národních jmenných prostorů.

¹⁸ Tj. znakový řetězec následující za „URN:NBN:“

¹⁹ <https://www.iso.org/obp/ui/>

²⁰ Konkrétní URN:NBN má pak například tuto podobu: urn:nbn:fi-fe20042357 .

²¹ Konkrétní URN:NBN má pak například tuto podobu: urn:nbn:fi:jyu-2011032410528 .

²² A dané dva syntaktické znaky, a to podle předepsaného modelu, tj. dvojtečku pouze pro členění národního jmenného prostoru na podprostory (identifikované kódem registrátora) a spojovník pro oddělení prefixu od řetězce NBN, a volitelně též pro případné dělení řetězce NBN.

Výzkum systémů URN:NBN

Užití definice

Pro potřeby výzkumu jsou užity následující termíny. *Digitálním dokumentem* se rozumí objekt identifikace, tedy to, co je identifikátorem označeno (tímto objektem může být jeden soubor, nebo skupina souborů). *Dlouhodobým digitálním úložištěm* se myslí archiv pro dlouhodobé uchovávání digitálních dokumentů v souladu s normou ISO 14721. *Digitální instancí* se rozumí jedna kopie digitálního dokumentu, která je umístěna na internetu za účelem zpřístupnění dokumentu uživatelům. Digitální instance musí mít logicky vždy určitou URL adresu, která identifikuje umístění této kopie a na základě které k ní přistupuje uživatel. *Aktuální URL adresou* se rozumí URL adresa, na níž je digitální dokument aktuálně zpřístupňován (URL adresy se mohou měnit). *Registrátorem* se rozumí subjekt (obvykle jde o instituci), kterému je přidělován identifikátor URN:NBN, registrační agenturou instituce zodpovědná za řízení tohoto přidělování.

Sledované kategorie

Pro potřeby sledování a srovnávání identifikačních systémů byly zvoleny následující kategorie: účastníci systému, objekt identifikace, model pro syntax, pravidla a technická infrastruktura.

Účastníci jsou vždy dvojího typu – registrační agentura a registrátoři. Důležité jsou zejména informace o tom, jaké typy institucí se systému účastní, jaké jsou jejich role a jaký je vztah této účasti s jinými aktivitami (např. dlouhodobým uchováváním).

V kategorii objekt identifikace sledujeme vymezení toho, čemu všemu lze v dané zemi přidělit identifikátor URN:NBN. Objekt identifikace může být vymezen například podle typu dokumentu nebo úrovně granularity (číslo, ročník periodika apod.).

Základní model pro syntax identifikátoru URN:NBN je stanoven standardem RFC 3188.²³ Ten uvádí jeho dvě možné varianty (se jmennými podprostory, nebo bez nich). Sledovaná kategorie si všímá toho, jaká varianta je užita a jakým způsobem jsou dané varianty v dané zemi dále konkretizovány nebo zpřísněny (jednotlivé národní knihovny si mohou zavést různá další omezení na syntax, přirozeně pouze v mezích základního modelu daného standardem RFC 3188).

Pravidla systému jsou nejobtížněji definovatelnou položkou, jelikož jsou v různých národních implementacích vymezena různým způsobem a v různé míře podrobnosti.

Pro analýzu a srovnání technické infrastruktury identifikačního systému byla zavedena následující dekompozice. Identifikační systém je vždy do určité míry řízen registrační agenturou. Na provozu jednotlivých prvků systému se mohou podílet další subjekty. Jako základní prvky infrastruktury, bez nichž se neobejde žádný národní systém URN:NBN, byly identifikovány následující komponenty: resolver, registr, generátor a evidenční služba.

Resolver je komponenta zajišťující čistě dereferenční službu. Dereference může znamenat přesměrování na aktuální URL adresu digitální instance, poskytnutí metadat k identifikovanému dokumentu nebo jiných informací. Resolver musí být vždy spojen s registrem.

Registr je komponenta, která musí obsahovat informace potřebné k dereferenci identifikátorů. Obvykle je tvořen databází vazeb identifikátorů URN:NBN a URL digitálních distancí.

²³ <http://www.rfc-base.org/rfc-3188.html>

Generátor je nástroj pro přidělování (přiřazování) identifikátorů URN:NBN, který zajišťuje jejich jedinečnost. Generátor však nemusí být vždy součástí identifikačního systému (tak tomu je v případě, kdy si registrátoři přidělují identifikátory URN:NBN samy užitím vlastních nástrojů).

Evidenční služba zajišťuje v první řadě registrování aktuálních URL adres, případně dalších metadat. Může být realizována například prostřednictvím sklízecího nástroje (harvester).

Kromě těchto výše uvedených komponent může být identifikační systém tvořen dalšími, zde neuvedenými. Stejně tak jsou pro chod identifikačního systému využívány další prvky, které nejsou vlastní součástí systému, ale komunikují s ním. Sem patří například informační systémy na straně registrátora, například OAI-PMH repository, v němž jsou vystavena metadata, která sklízí evidenční služba identifikačního systému.

Metoda výzkumu

Jako základní metoda výzkumu byla použita analýza oficiálních informací, dostupných na webových stránkách registračních agentur (národních knihoven) sledovaných zemí, a publikovaných článků a příspěvků od odborníků, kteří se na budování daných identifikačních systémů podílejí (viz odkazy v textu). Jelikož se, jak se předpokládalo, ukázaly tyto informace jako nedostatečné, bylo využito též emailové komunikace se zástupci daných systémů a v jednom případě i osobní komunikace (služební cesta do Německa), a to za účelem zjištění podrobnějších informací nebo zpřesnění či odstranění nejasností.

Dále byla využita systémová analýza komponent technických infrastruktur jednotlivých identifikačních systémů, kterou provedl Martin Řehánek v rámci interní studie [Řehánek, 2015]. V rámci této analýzy byly testovány zejména jednotlivé resolvers a generátory, včetně instalace komponent tam, kde to bylo možné (kde byly k dispozici online nástroje).

Zdůvodnění výběru

Identifikační systémy využívající identifikátory URN:NBN byly zavedeny v několika evropských zemích (především ve Skandinávii a ve střední Evropě). Standard pro užívání identifikátoru dovoluje poměrně široké možnosti nastavení národních implementací. Neklade žádné nároky na technické provedení a i pokud se jedná o organizační pravidla, ponechává značnou benevolenci. V důsledku těchto skutečností vznikly různé identifikační systémy s odlišným technickým řešením i pravidly přidělování a správy. Systémy se liší i úrovní poskytovaných služeb i celkovou architekturou systému a podobou informací, které poskytují uživatelům. V následujících částech budeme analyzovat vybrané národní implementace identifikačních systémů založených na identifikátorech URN:NBN.

Pro analýzu byly zvoleny systémy Německa, Finska, Nizozemsko a Itálie. Každý z těchto systémů je v jistém ohledu specifický a z hlediska implementace standardu URN:NBN významný. Jako celek vhodně ilustrují možnosti práce s identifikátorem URN:NBN.

Finský systém je hodný pozornosti již tím, že v této zemi byl identifikátor URN:NBN navržen a jako první využíván. Na druhou stranu počet aktuálně identifikovaných dokumentů řadí finský systém spíše k těm menším. Německá implementace naopak představuje největší systém využívající URN:NBN, službu poskytuje národní knihovna a zabezpečuje i požadavky Rakouska a Švýcarska na správu identifikátorů. I Nizozemsko se řadí spíše k větším uživatelům identifikátoru URN:NBN, jeho systém je zvláštní podvojností v organizaci využívání služeb. Zároveň nizozemské knihovny spravují jednu z největších sbírek e-born vědeckých publikací. Mj. i díky tomu, že mnohá prestižní nakladatelství vědecké literatury sídlí právě v Nizozemsku. Zajímavé možnosti nabízí italská implementace. Je to dáno jednak jejím spojením s povinným výtiskem elektronických publikací a také originálním řešením

podpůrného softwaru. Důležité je zmínit i to, že italský systém patří k nejstarším návrhům, jeho první verze vznikla již v roce 2007.

Analýza některých národních implementací v rámci systému URN:NBN

Finsko

Za zrodem standardu pro URN:NBN stojí Juha Hakala z finské národní knihovny (Kansalliskirjasto; dále jen jako „NLF“²⁴). Ostatně NLF se aktivně podílela na standardizaci identifikátorů URN:NBN již od 90. let 20. století. NLF zprovoznila svůj národní resolver v roce 2007, ale přidělování identifikátorů URN:NBN dokumentům začalo o několik roků dříve, kdy NLF umožnila, aby se jiné instituce staly registrátory finského systému a přidělovaly identifikátory URN:NBN pro své vlastní digitální dokumenty. [Keskitalo, 2015; Řehánek, 2012; Voigt, 2007]

Účastníci

Pro Finsko vykonává roli registrační agentury NLF. Ta provozuje identifikační systém, sestávající z resolveru, registru, generátoru a sklízecího nástroje. Registrátorem finského systému se mohou stát pouze instituce (jednotlivci nikoliv), v současnosti jich je zapojeno zhruba 50, převážně univerzit. [Keskitalo, 2015; Voigt, 2007]

Objekt identifikace

Identifikátor URN:NBN je povoleno přidělit pouze dokumentům, které jsou online dostupné na webové stránce. Omezení na typ obsahu dokumentu není zavedeno (může jít o textový, obrazový i zvukový obsah). Identifikátor URN:NBN lze dokonce přidělit i sbírce dokumentů. [Kansalliskirjasto, 2007] V praxi jsou identifikátory URN:NBN nejčastěji přidělovány univerzitním online publikacím a výzkumným datům. [Keskitalo, 2015] Dosud bylo evidováno přes 300 tisíc identifikátorů URN:NBN.²⁵

Model pro syntax

Ve finském systému se vyskytují následující modely pro syntax: [Řehánek, 2015]

1.	urn:nbn:fi-<řetězec NBN>	
1.1	řetězec NBN:	fe<rok přidělení><číslo>
1.2.	řetězec NBN:	fe<rok přidělení><měsíc přidělení><den přidělení><číslo>
2.	urn:nbn:fi:<kód registrátora>-<řetězec NBN>	
2.1	řetězec NBN:	<rok přidělení><měsíc přidělení><den přidělení><číslo>
2.2	řetězec NBN:	<alfanumerický řetězec>

Schéma řetězce NBN je ve finském systému dále konkretizováno následujícím způsobem.

V prvním modelu (s nečleněným národním prostorem) jsou uvedena schémata řetězců NBN daná generátorem (1.2 pro současnou verzi generátoru, 1.1 pro dřívější).

Ve druhém modelu (se jmennými podprostory) je pod číslem 2.1 uvedeno schéma řetězce NBN pro vybrané čtyři registrátory, přednastavené v generátoru.²⁶ Tito čtyři registrátoři si mohou generovat řetězec NBN tímto nástrojem. Není to však nutností, pokud nechtějí,

²⁴ Zkratka z oficiálního anglického názvu „The National Library of Finland“ (<http://www.kansalliskirjasto.fi/>).

²⁵ Počet přidělených URN:NBN může být samozřejmě přirozeně vyšší, vzhledem ke způsobu přidělování.

²⁶ Vědecká knihovna Tilastotieteiden kirjasto (kód "libstat"), univerzitní knihovny Jyväskylän yliopiston kirjasto ("jyu") a Lapin yliopiston kirjasto ("ula") a výzkumný ústav Metsäntutkimuslaitos ("metla")

mohou si generovat řetězce NBN samy (v rámci svého podprostoru), a pak pro ně, jako pro všechny ostatní registrátory, platí model 2.2, kde nemusí být užito pouze čísel.

Další varianty pro syntax (zmíněné v samotném standardu pro URN:NBN) se reálně nevyskytují.²⁷

Pravidla

Základním pravidlem je, že registrátoři musí nést zodpovědnost za správné přidělování identifikátorů URN:NBN, zajistit trvalou údržbu online přístupu k identifikovaným dokumentům a umožnit využívání identifikátorů URN:NBN k přesměrovávání na aktuální lokaci dokumentu na webu.

Pro přidělování platí následující pravidlo. Pokud digitálnímu dokumentu lze přidělit identifikátor ISBN, pak tomuto dokumentu musí být nejprve přiděleno číslo ISBN a z něho vytvořen identifikátor. [Finsko, 2014] Ten je tvořen jednoduše podle modelu „URN:ISBN:<číslo ISBN>“. Například online kniha s číslem ISBN 952-10-0093-7 má identifikátor URN:ISBN:952-10-0093-7. Trvalá URL adresa je vytvářena podle schématu <adresa resolveru><identifikátor URN:ISBN>. Například uvedená kniha bude mít URL <http://urn.fi/URN:ISBN:952-10-0093-7>. To nás nyní přesměrovává na aktuální URL adresu <https://helda.helsinki.fi/handle/10138/18338>. Systém URN:ISBN je spravován finskou národní agenturou pro ISBN.

Identifikátor URN:NBN lze přidělit pouze tehdy, pokud dokument nemá nebo mu nelze přidělit číslo ISBN. Registrační agentura doporučuje vytvářet trvalé URL odkazy na dokumenty podle schématu <adresa resolveru><identifikátor URN:NBN>. [Finsko, 2014]

Identifikátor URN:NBN musí být zaznamenán do metadat digitálního dokumentu. Pokud je to možné, musí být též zobrazen uvnitř dokumentu samotného.²⁸ Identifikátor URN:NBN lze zapisovat malými i velkými písmeny (resolver není citlivý na velikost písma).

Přidělování samotné je v odpovědnosti registrátorů. Ti si mohou generovat řetězce NBN sami vlastními prostředky (v tomto případě dostanou přidělený podprostor). V tomto případě ručí za jedinečnost řetězce NBN. Například Aalto University²⁹ má přidělen kód „aalto“ (příkladem identifikátoru je pak URN:NBN:fi:aalto-201305166317).

Druhou možností je nechat si přidělit identifikátor URN:NBN online generátorem, který provozuje NLF (v tomto případě však bude identifikátor obsahovat samotný národní prostor; vyjma čtyř vybraných registrátorů). [Keskitalo, 2015]

Identifikátor URN:NBN je však závazný až po připojení na URL digitální instance v registru identifikačního systému. Tento proces probíhá ve dvou krocích. Nejprve si registrátor vygeneruje identifikátor URN:NBN a zaznamená jej do svého dokumentu, který musí být dostupný online. Následně registrátor umožní identifikačnímu systému evidovat všechny své přidělené identifikátory URN:NBN, přičemž ke každému identifikátoru URN:NBN musejí dodat nejméně jednu funkční URL adresu digitální instance (případně alespoň informaci, že digitální dokument již není dostupný). URL adresa může odkazovat buď přímo k digitálnímu dokumentu, nebo k jeho bibliografickému záznamu tohoto, případně ke vstupní stránce (URL adrese, ze které je dokument dostupný). Právě vstupní stránka by měla být dostupná všem uživatelům v případě, že přístup k dokumentu je omezen. [Keskitalo, 2015; Pietarilla, 2015]

²⁷ urn:nbn:fi-fea-5c5875e6e49ae649cad63e5ee4f6c346, urn:nbn:fi:st:, urn:nbn:fi:vn

²⁸ V praxi jsme však zaznamenali řadu případů, kdy toto pravidlo není vždy dodržováno. Viz například dokument s identifikátorem URN:NBN:fi-fe20042357, v jehož těle tento identifikátor uveden není, ačkoliv jej do něj zapsat možné je (jde o soubor ve formátu PDF).

²⁹ <http://www.aalto.fi/en/>

Všechny identické kopie digitálního dokumentu obdrží tentýž identifikátor URN:NBN. Dereference jednoho URN:NBN tudíž může zahrnovat přesměrování na několik odlišných URL adres, pořadí určuje nastavení systému. Pokud je obsah digitálního dokumentu signifikantně změněn, je nutno přidělit nový identifikátor URN:NBN. V případě drobných změn toto vyžadováno není.

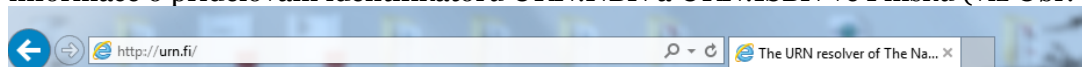
Těm institucím, které jsou schopné zajistit, aby si od nich identifikační systém mohl automaticky sklízet aktuální URL adresy digitálních instancí, je přidělen samostatný jmenný podprostor v rámci finského národního prostoru. Ostatní registrátoři si musí generovat identifikátory URN:NBN v rámci národního prostoru jako takového (URN:NBN:fi). [Keskitalo, 2015]

Na vstupní stránce, ze které registrátor zpřístupňuje identifikovaný dokument, by měla být uvedena následující metadata: trvalý odkaz, název dokumentu, autora, v případě vysokoškolské práce její typ, abstrakt, datum a přístupová práva. K dispozici jsou případně i soubory (tvořící identifikovaný dokument) ke stažení či náhledu s dostupnými informacemi o velikosti a typu souboru.

Technická infrastruktura

Finský systém je založen na jediném, celonárodním **resolveru**, který je provozován NLF. Softwarové řešení bylo vyvinuto samotnými zaměstnanci NLF. Resolver dokáže vykonat dereferenci nejen pro finské identifikátory URN:NBN, ale též pro identifikátory URN:ISBN přidělené finským knihám dostupným online. [Pietarilla, 2015]

Samotný resolver je dostupný na adrese <http://urn.fi/>. Jeho webové rozhraní je minimalistické. Jsou zde pouze dva řádky textu, v druhém řádku jsou odkazy na základní informace o přidělování identifikátorů URN:NBN a URN:ISBN ve Finsku (viz Obr. 1).



This is the [URN](#) resolver of [The National Library of Finland](#).

More information is available in [English](#), [Finnish](#) and [Swedish](#).

Obr. 1 - Webové rozhraní finského resolveru

Finský resolver umí vykonávat pouze standardní přesměrovávání na URL digitální instance. Identifikátor URN:NBN nelze zadat nikam do webového rozhraní resolveru, přesměrovávání funguje prostým užitím URL adresy ve tvaru <adresa resolveru><identifikátor URN:NBN>. [Řehánek, 2015]

K přidělování identifikátorů URN:NBN může registrátor využít online **generátor**, který provozuje sama NLF. Generátor je dostupný na webovém sídle Helsinské univerzity

(<http://urnsource-kk.lib.helsinki.fi/cgi-bin/urn.cgi>).³⁰ Jedná se o jednoduchou webovou aplikaci napsanou v Pythonu. Tento nástroj lze využít dvěma způsoby. Prvním možností je, že si registrátor přes webové rozhraní manuálně vygeneruje 1-10 identifikátorů URN:NBN v jedné dávce (viz Obr. 2). Druhou možností je, že se registrátor připojí svým informačním systémem k tomuto generátoru prostřednictvím rozhraní API.

V případě užití generátoru je identifikátor URN:NBN vytvořen přímo v národním prostoru Finska. Výjimku tvoří 4 vybraní registrátoři, kteří si pomocí tohoto generátoru mohou nechat přidělit identifikátor URN:NBN v rámci svého vlastního jmenného podprostoru, který je v generátoru již přednastaven (nemusí tak však činit a mohou si generovat identifikátory sami, jako jakýkoliv jiný registrátor s přiděleným vlastním podprostorem). Nicméně pouze v případě jednoho jmenného podprostoru³¹ nám nástroj neumožnil nechat si vygenerovat identifikátor URN:NBN. Autorizace tedy pravděpodobně probíhá pouze na úrovni IP adresy, která je navíc buď špatně konfigurovaná, nebo vypnutá. Jediné omezení pro generátor je denní počet vygenerovaných identifikátorů z jedné IP adresy.³² Jako výchozí nastavení je vybrán celý finský jmenný prostor.³³ [Řehánek, 2015]

URN-tunnusten luontiohjelman

Ohjelma rakentaa kansallisbibliografian ID-tunnukseen perustuvia URN-tunnuksia. Tunnusten syntaksi on sovittu Internet Engineering Task Force'n (IETF) URN-työryhmän kanssa. Tunnusten pysyvyyden ja ainutkertaisuuden takaa Kansalliskirjasto. Generaattoria voi käyttää myös koneellisesti ja voit varata oman URN-alinimiavaruuden Kansalliskirjastolta, jos siihen on tarvetta. Ohjeet generaattorin käyttöön löytyy [täältä](#). Lisätietoja [URN-opasteestamme](#) ja osoitteesta: [kk-urn AT helsinki PISTE fi](mailto:kk-urn@helsinki.fi)

Tunnusten tyyppi	Maa	Nimiavaruuden osa	URN-tunnusten määrä	
NBN	Suomi	Koko suomalainen nimiavaruus	1	Hae URN-tunnus

2
3
4
5
6
7
8
9
10

Obr. 2 - Nástroj pro manuální generování URN:NBN ve správě NLF (17. 6. 2015)

Generátor není spojen s registrem identifikačního systému. Současná verze generátoru je robustnější než předchozí v tom smyslu, že je nutné ukládat vygenerované identifikátory (resp. čítače pro jednotlivé registrátory) pouze pro aktuální den (a ne pro celý rok, jak tomu bylo v předchozí verzi), takže jejich případná ztráta není zcela fatální pro celý systém. Nicméně tento způsob generování představuje riziko vzniku určitého (byť výrazně menšího) počtu nejedinečných identifikátorů URN:NBN.³⁴

Evidenční služba probíhá následovně. Vzhledem k tomu, že generátor není spojen s registrem resolveru, musejí registrátoři umožnit evidenci nejen aktuálních URL adres digitálních instancí, ale spolu s nimi (při první evidenci) též samotných identifikátorů URN:NBN do registru resolveru. Ve finském systému evidenční službu technicky vykonává **sklízecí nástroj**, napsaný v Pythonu. Sklizení přidělených identifikátorů URN:NBN a URL

³⁰ NLF je oficiálně součástí Helsinské univerzity.

³¹ Kód „libstat“

³² http://urnsource-kk.lib.helsinki.fi/fi_kk_urn_generaattori_ohje.pdf

³³ Finsky „koko suomalainen nimiavaruus“.

³⁴ Rozumným řešením případné ztráty by bylo okamžité vypnutí generátoru po zbytek dne, kdy se čítače ztratily.

adres digitálních instancí u registrátora probíhá minimálně jednou denně a může se odehrávat dvěma způsoby. [Pietarilla, 2015; Řehánek, 2015]

Prvním mechanismem evidenční služby je klasické sklizení metadat přes protokol OAI-PMH. Jsou sklizeny všechny záznamy ve formátu Dublin Core (dc-xml) v systému registrátora. Pokud nástroj v záznamu objeví identifikátor URN:NBN a URL adresu (v polích dc.identifier), ve své databázi si vytvoří (nebo aktualizuje) záznam obsahující tuto vazbu.

Pokud systém registrátora protokol OAI-PMH nepodporuje, je druhou možností poskytnout NLF statickou URL adresu, na které registrátor vystaví jeden (velký) soubor v XML obsahující všechny vazby URN:NBN/URL. Sklízecí nástroj pak takto vystavený soubor v XML pravidelně stahuje, prochází jednotlivé záznamy a do vlastní databáze si ukládá tyto vazby. Konkrétní XML formát vychází ze švédského systému, který byl dříve používán i v NLF.³⁵ Záznamy z databáze sklízecího nástroje si pravidelně přebírá registr resolveru (založený, stejně jako databáze sklízecího nástroje, na dokumentové databázi CouchDB), který na základě jejich zpracování provádí svou aktualizaci.

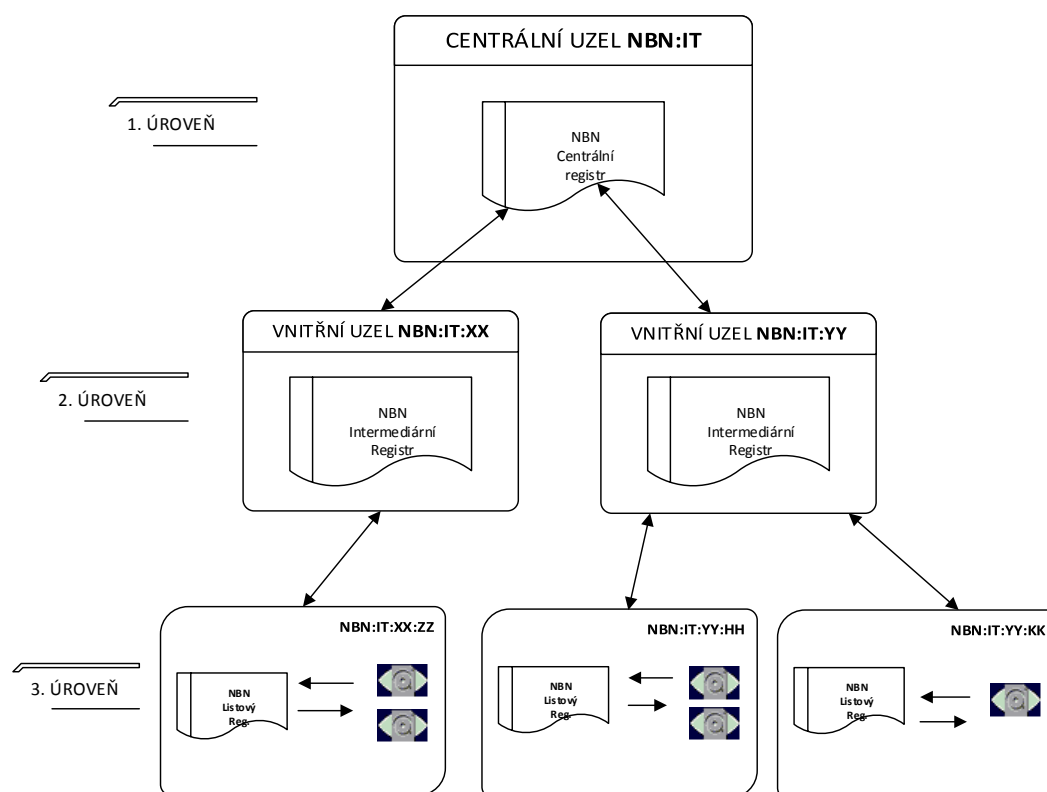
Sklizení v obou případech probíhá několikrát denně. V současné době musí mít NLF plnou důvěru v organizace, že udržují dokumenty online dostupné. [Keskitalo, 2015]

Itálie

Vývoj italského identifikačního systému je úzce svázán se zdejším systémem povinného výtisku (Zákon 106/2004; Decreto del presidente della repubblica 252/2006),³⁶ který pokrývá i digitální dokumenty. V souvislosti s ním začaly od roku 2007 práce na vybudování italského identifikačního systému. Ty se následně odehrávaly ve třech etapách. V každé z nich se měnil model celého systému i softwarové technologie. (V první etapě Národní knihovna ČR zvažovala využít toto softwarové řešení, ale z důvodu nízké podpory a odlišného modelu bylo nakonec od takové varianty ustoupeno). V druhé etapě byl představen níže zobrazený třístupňový hierarchický model systému (viz Obr. 3).

³⁵ Konkrétní XML schéma je dostupné online (<http://epc.uu.se/schema/rs/3.0/rs-location-mapping-schema.xsd>)

³⁶ http://bibliotecaestense.beniculturali.it/docvar/i-mo-beu_reg-dep-dpr-2006-252.pdf



Obr. 3 - Původní třístupňová hierarchická struktura italského systému

Centrální uzel byl zodpovědný za správu národního prostoru (vytváření, přemístění nebo odstranění subdomén; práva k přidělování identifikátorů aj.). Delegoval zodpovědnost na domény druhé úrovně (registrační agentury druhé úrovně zodpovědné za přiřazování jmen institucím na úrovni pod nimi). Vespuďu byly tzv. listové uzly, jejichž úkolem bylo sklízet metadata z úložišť a přidělovat digitálním dokumentům jedinečné identifikátory. [Bellini, 2010]

V současné etapě bylo od třístupňového modelu odstoupeno a byl spuštěn dvojstupňový model (popsaný dále).

Společně s identifikačním systémem NBN:IT je budováno dlouhodobé digitální úložiště, nazvané Magazzini Digitali (MD)³⁷, určené zejména pro digitální publikace získané z povinného nebo dobrovolného výtisku. [Bellini, 2014] Formátem balíčku AIP je zde WARC (tedy kontejnerový formát užívaný ve webarchivech). Do úložiště MD dodávají obsah (získaný z povinného nebo dobrovolného výtisku) tři knihovny: Biblioteca Nazionale Centrale di Firenze (BNCF, <http://www.bncf.firenze.sbn.it/>), Biblioteca Nazionale Centrale di Roma (BNCR, <http://www.bncrm.librari.beniculturali.it/>) a Biblioteca Nazionale Marciana di Venezia (<http://marciana.venezia.sbn.it/>). Na systému se dále podílejí instituce Fondazione Rinascimento Digitale (FRD, <http://www.arcamemorie.net/>) a Konference rektorů italských univerzit (Conferenza dei Rettori delle Università Italiane, CRUI, <http://www.cruai.it/>). Konference dohlíží na testování ukládání závěrečných vysokoškolských prací.

Je třeba uvést, že Italové jednak prošli zřejmě nejsložitějším vývojem softwarového řešení pro identifikátory URN:NBN, jednak patří také mezi nejvíce publikující odborníky v oblasti trvalých identifikátorů.

Tvůrci italského systému rozlišují tři roviny perzistence. [Bellini, 2012] Perzistence identifikátoru je závazek zachovat identifikátor (i když referent zanikl). Uchováváním

³⁷ Česky digitální úložiště.

metadat v registru lze ověřit minulou existenci referentu. Perzistence vazby URN:NBN/URL je závazek, že lze provést základní dereferenci (i když již dokument není online dostupný). Třetím typem perzistence je perzistence označovaného zdroje. Ta sice formálně není záležitostí vlastního identifikačního systému (jde o funkci dlouhodobého digitálního úložiště), nicméně italský systém je právě úzce spojen s MD, do nějž jsou ukládány kopie identifikovaných dokumentů. Tím je - alespoň teoreticky - v případě, že poskytovatel obsahu přestane dokument zpřístupňovat, zaručena možnost dereferencí získat přístup k označovanému objektu.

Účastníci

Registrační agenturou italského identifikačního systému je BNCF, jedna z dvojice italských národních knihoven.

Registrátory jsou v italském systému instituce, které ukládají své digitální dokumenty do úložiště MD.³⁸ Jedná se zejména o univerzity. Platí, že pouze to, co je uloženo do MD, může získat identifikátor URN:NBN. [Bellini, 2012]

Objekt identifikace

Identifikátor URN:NBN lze přidělit pouze dokumentům, jejichž kopie jsou určeny pro dlouhodobé uchování v úložišti MD v rámci povinného nebo dobrovolného elektronického výtisku. V současnosti se jedná o elektronické publikace univerzit a v testovací fázi je přidělování identifikátorů URN:NBN dizertačním pracím (tesi di dottorato). (Na dizertační práce se vztahuje zákon.)³⁹ Dále se plánuje identifikovat digitalizáty vzniklé v knihovnách i jiných institucích. Zcela na dobrovolné bázi lze ukládat také elektronické knihy vydavatelů.

Podle pravidel odpovídá objekt identifikace úrovni reprezentace ve standardu PREMIS.⁴⁰

Dosud bylo v Itálii přiděleno zhruba 15 tisíc identifikátorů URN:NBN.⁴¹

Model pro syntax

V současnosti platí v italském systému jeden jednotný model pro syntax pro všechny případy: [Řehánek, 2015]

1.	urn:nbn:it:<kód registrátora>-<řetězec NBN>
----	---

Kódem registrátora je zpravidla zkratka dané instituce. Řetězec NBN obsahuje v současných případech pouze čísla. Například Univerzita ve Florencii (Università degli Studi di Firenze, UNIFI, <http://www.unifi.it/>) obdržela jmenný podprostor s kódem „unifi“, příkladem je identifikátor urn:nbn:it:unifi-3866. Záměrem tvůrců italského systému je, aby řetězec NBN nebyl významotvorný.

Pravidla

URN:NBN lze přidělit pouze dokumentu, který bude uložen do MD.

³⁸ Viz http://nbn.depositolegale.it/nbn_report.html

³⁹ <http://www.depositolegale.it/deposito-legale-digitale-delle-tesi-di-dottorato/>

⁴⁰ Podle standardu PREMIS je reprezentace taková množina počítačových souborů (včetně strukturálních metadat, např. uložených v textovém souboru v textovém souboru v XML), která je nutná k tomu, aby mohla být softwarovou aplikací plně a smysluplně reprodukována jako jedna intelektuální entita (intelektuální entita je množina informací, která je považována za jednu intelektuální jednotku z hlediska správy a popisu, tedy například kniha na úrovni titulu).

⁴¹ http://nbn.depositolegale.it/nbn_report.html

BNCF jako registrační agentura přiděluje registrátorům celý identifikátor URN:NBN centrálně, a to výměnou za dodání požadovaných metadat (pouze URL adresa digitální instance a URL adresa metadatové stránky popisující identifikovaný dokument, ideálně přístupné přes protokol OAI-PMH). BNCF poté sklídí z dodaných URL adres metadata i digitální dokument samotný a společně je uloží do MD. Po uložení je vytvořena nová URL adresa odkazující přímo do digitálního úložiště (dokument je z digitálního úložiště zpřístupňován systémem WayBack Machine oprávněným uživatelům). V tento moment se stává identifikátor URN:NBN aktivním (tj. resolver může vykonávat jeho dereferenci). Přidělení identifikátoru URN:NBN je v italském systému tedy propojeno s klasickým dodáváním dokumentu (zde formou sklizení webu) do digitálního úložiště k dlouhodobému uchovávání.

V registru systému tedy nejsou udržována bibliografická metadata o referentech, pouze administrativní metadata popisující události s identifikátory v samotném registru. Jsou však uchovávány odkazy na tato metadata uložená v digitálním úložišti.

Při evidenci registrátora do italského systému probíhá následující proces:

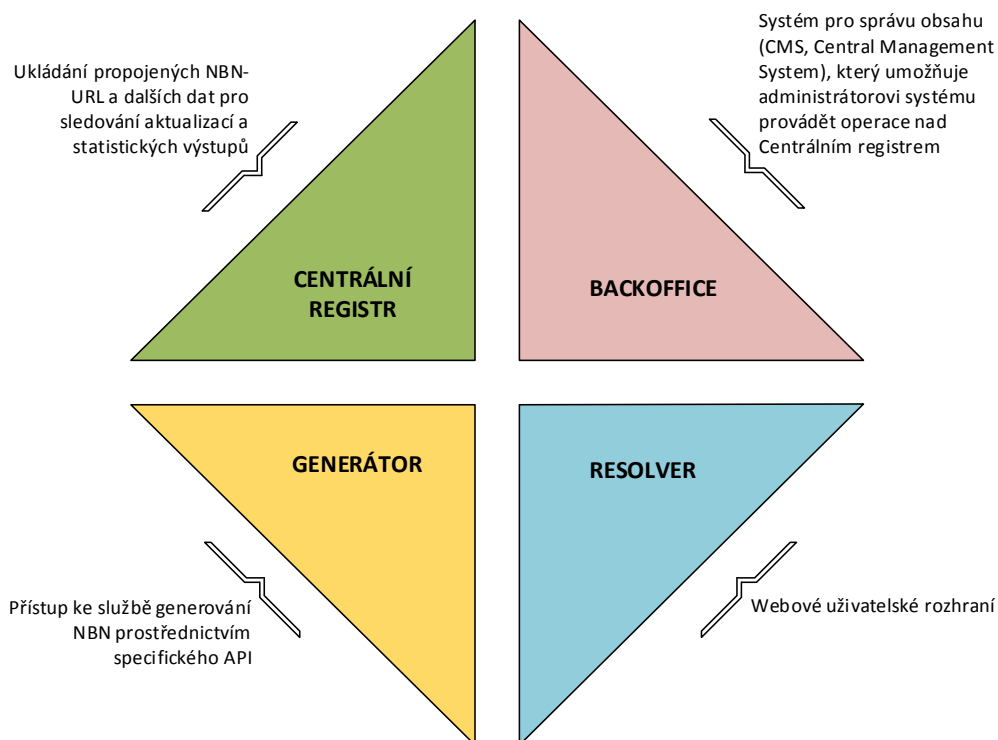
- Předběžné hodnocení obsahu, který má být uložen do MD.
- Smlouva mezi BNCF a potenciálním registrátorem.
- Případná konzultace problematiky ochrany práv duševního vlastnictví a také licenční vymezení pro sbírky.
- Registrace institucionálního repozitáře⁴² (zastoupen odpovědným agentem-uživatelé s přihlašovacím jménem a heslem)
- Přiřazení jmenného podprostoru. [Bellini, 2012]

Technická infrastruktura

Technická infrastruktura italského identifikačního systému se skládá ze čtyř komponent, které jsou vzájemně propojeny: generátor, centrální registr, systém pro správu obsahu (pro správu registru) a resolver (viz Obr. 4). Dvoustupňový model je založen na identifikačním systému na straně jedné a pluginu na straně registrátora.⁴³ [Piščans, 2013]

⁴² Tedy informačního systému registrátora, ve kterém zpřístupňuje své dokumenty

⁴³ Pluginy zajišťují komunikaci s registrační službou přes RESTful API. Jsou vydané pod open source licencemi a jejich zdrojové kódy jsou dostupné na githubu: 1) EPrints plugin - <https://github.com/depositolegale/eprints-nbn>; 2) DSpace plugin - <https://github.com/depositolegale/dspace-nbn>; 3) OJS plugin - <https://github.com/depositolegale/ojs-nbn>



Obr. 4 – Komponenty italského systému

Centrální registr obsahuje záznamy všech vazeb URN:NBN/URL. Jsou zde zpravidla tři typy URL adres vázaných k jednomu identifikátoru URN:NBN: adresa dokumentu u registrátora, adresa na popisná metadata a adresa dokumentu uloženého v MD.

Registr uchovává administrativní informace (metadata o veškerých událostech, které proběhnou nad identifikátorem), a to s cílem „sledovat status záznamu a zaznamenávat životní cyklus identifikátoru URN:NBN podle filozofie PREMIS.“ [Bellini, 2014] (viz Tab. 1)

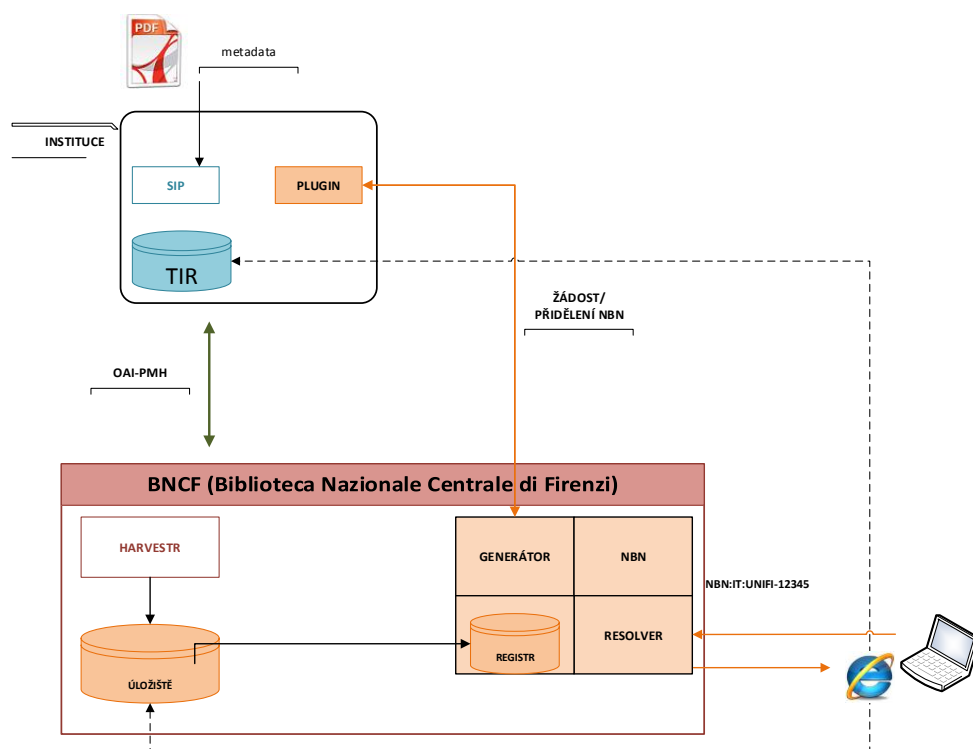
POLE	HODNOTA
Status	<i>Temporary</i> : dokument ještě nebyl sklizen z digitálního úložiště <i>Stable</i> : dokument je uložen v digitálním úložišti
Provozní informace	<i>Active</i> : přesměrování URN:NBN je možné <i>Off</i> : přesměrovávání URN:NBN není možné (je pozastavené)
Akce	<i>NBN_Create</i> : vytvoření nového URN:NBN <i>NBN_Update</i> : nové propojení (aktualizace) vazby URL/URN:NBN <i>Status_update</i> : status URN:NBN se změnil

Tab. 1 - Sledování životního cyklu identifikátoru URN:NBN

V italském systému je v současnosti pouze jeden celonárodní **resolver**, jehož webové rozhraní je dostupné na <http://nbn.depositolegale.it/>. Resolver provádí přesměrování na webovou stránku obsahující administrativní metadata o identifikátoru URN:NBN (status URN:NBN a URL adresy), dále odkaz na stránku poskytovatele obsahu; odkaz na kopii dokumentu uloženou v MD a odkaz na metadata uložená tamtéž (a zpřístupňovaná prostřednictvím WayBack Machine).

Generátor je propojen s registrem systému, takže resolver okamžitě ví o přidělených identifikátorech URN:NBN. Čísla v řetězci NBN jsou generována vzestupně pro všechny registrátory, společné počítadlo generuje postupně se zvyšující čísla dosazována jako ID dokumentu.⁴⁴ Přidělování identifikátorů URN:NBN se v italském systému uskutečňuje výlučně tímto generátorem, a to na základě přihlášení registrátora.

Evidenční služba probíhá ve dvou krocích. Nejprve identifikační systém přidělí registrátorovi URN:NBN na základě dodaných metadat. Tento proces se odehrává komunikací mezi informačním systémem registrátora a generátorem identifikačního systému, a to prostřednictvím RESTful API rozhraní (založeném na JSON protokolu).⁴⁵ Registrátor musí mít v identifikačním systému založen přihlašovací účet, komunikace je autorizována technologií Digest access authentication⁴⁶ (viz Obr. 5). [Bellini, 2012]



Obr. 5 - Komunikace mezi registrátorem a identifikačním systémem

Registrátor zašle požadovaná metadata⁴⁷ a identifikační systém okamžitě vrátí přidělený identifikátor URN:NBN ve formátu JSON. Do databáze identifikačního systému jsou vloženy přidělený identifikátor URN:NBN a obě URL adresy. Stav identifikátoru URN:NBN je zatím označen jako přechodný (temporary) a resolver zatím neprovádí přesměrovávání. Druhou variantou prvního kroku je, že si registrátor nainstaluje do svého systému pro správu obsahu plugin, který mu umožní přidělit identifikátor URN:NBN již při samotném vydávání publikace. V současnosti je plugin dostupný pro systémy EPrints, OJS (Open Journal Systems, <http://pkp.sfu.ca/ojs/>) a DSpace. Plugin umožňuje dialog s identifikačním systémem přes webovou službu. Ušetří se tím čas a úsilí při vkládání identifikátoru až po odevzdání povinného výtisku a zároveň se usnadní využití identifikátoru, který je přítomen již při prvním uveřejnění. [Bellini, 2014; Bellini, 2015]

⁴⁴ Tj. řetězec „3866“ v urn:nbn:it:unifi-3866 neznamená 3866. identifikátor v rámci UNIFI.

⁴⁵ http://nbn.depositolegale.it/api/nbn_generator.pl

⁴⁶ <https://tools.ietf.org/html/rfc2617>

⁴⁷ Parametry: action = nbn_create / url = [URL adresa digitální instance] / metadataURL = [URL adresa metadatového záznamu]

Druhým krokem evidence je, že identifikační systém provede sklizení metadat i samotného dokumentu a vytvoří novou URL stránku odkazující do MD. Na tuto novou stránku je pak prováděno přesměrovávání (jde tedy o úvodní stránku dokumentu). V tento okamžik je změněn status identifikátoru URN:NBN na aktivní (tj. přesměrovávání je funkční). Všechny tři odkazy (na metadatovou stránku, na dokument v MD uložený a na dokument uložený ve zpřístupňovacím systému registrátora) jsou evidovány v registru identifikačního systému, ze kterého resolver čerpá.

Nizozemsko

V Nizozemsku se rozhodli využít standard URN:NBN pro identifikátory digitálních dokumentů v roce 2007. Iniciativa vzešla z Královské nizozemské akademie věd (Koninklijke Nederlandse Akademie van Wetenschappen, KNAW, <https://www.knaw.nl/nl>) a Nizozemské organizace pro vědecký výzkum (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO, <http://www.nwo.nl/>), které navázaly kontakt s Královskou knihovnou (Koninklijke Bibliotheek, KB, <https://www.kb.nl/>), jež byla jakožto národní knihovna oprávněna systém URN:NBN zaštitit. Iniciativa byla vedena snahou zajistit trvalý přístup k digitálním dokumentům vědecké a výzkumné provenience, které zpřístupňoval portál narcis.nl. Zejména šlo o elektronické knihy (ty bylo obtížné identifikovat jinak), odborné publikace a výzkumná data. Zmíněný portál existoval již dříve a právě narůstající počet ukládaných dokumentů vedl k potřebě zavést nový identifikační systém. KNAW a NWO společně zřídili organizaci DANS (Data Archiving and Networked Services, <http://www.dans.knaw.nl/nl>), která má za úkol spravovat identifikační systém. KB se později rozhodla identifikátory URN:NBN využít i pro své dokumenty, což vedlo až k současnému stavu, kdy v této zemi existují dva resolvers.

Účastníci

Oficiálně spravuje národní identifikační systém založený na URN:NBN (dále jen „systém URN:NBN:NL“) Národní agentura pro NBN (Registration Agency NBN), která je součástí KB. KB jako registrátor má také vlastní resolver a specifickou syntax. Další skupinou jsou výzkumné a vědecké instituce, zejména výzkumné univerzity.

Objekt identifikace

KB jakožto registrátor využívá identifikátory pro své vlastní fondy.⁴⁸ Jedná se zejména o elektronické knihy komerčních vydavatelů (dobrovolný výtisk), vědecká periodika (na úrovni článku) a elektronické noviny a časopisy (na úrovni čísla). KB dosud nepřiděluje identifikátory URN:NBN svým digitalizátům. To se plánuje až po migraci digitalizovaného obsahu do nové infrastruktury dlouhodobého digitálního úložiště e-Depot. Identifikace elektronických knih a vědeckých článků je zatím prioritou.

Výzkumné instituce užívají URN:NBN převážně pro své odborné publikace a výzkumná data. (Ty užívají identifikátory tvořené podle níže uvedeného druhého modelu.)

Při identifikaci se nehledí na verze v různých formátech, například tentýž článek dostane stejné URN:NBN pro verzi v PDF i v HTML.⁴⁹

Model pro syntax

V současnosti v nizozemském systému existují dva modely pro syntax:

1.	urn:nbn:nl-<řetězec NBN>
2.	urn:nbn:nl:kb-<řetězec NBN>
3.	urn:nbn:nl:<kód typu registrátora>:<kód registrátora>-<řetězec NBN>

⁴⁸ Například urn:nbn:nl:kb-1382537057536.

⁴⁹ Např. Identifikátor URN:NBN:NL:UI:10-1-116866 identifikuje článek „Just in case, just in time, or just don't bother? Assessment of one-shot library instruction with follow-up workshops“, který je dostupný v PDF i HTML (<http://liber.library.uu.nl/index.php/lq/article/view/URN:NBN:NL:UI:10-1-116866>).

První dva modely jsou určeny pouze pro KB jakožto registrátora.⁵⁰ Reálně se užívá zatím pouze druhý. KB již přidělila více jak 23 milionů dokumentů. Číslo je vysoké vzhledem k enormní sbírce publikací vydavatelů, které KB již delší dobu přijímá.

Třetí varianta se jmennými podprostory je dokonce dvojestupňová. První úroveň je tvořena skupinovým kódem. V současnosti jsou zavedeny dvě skupiny: pod kódem „UI“ (určeným pro výzkumné instituce, zejména jde o univerzity) je nyní zařazeno 26 registrátorů a identifikováno bylo již více než 1,4 milionu dokumentů. Kód „HS“ je určen pro ostatní vysoké školy a další instituce vyššího vzdělávání, dosud je registrováno 24 institucí.

Pravidla

Registrátoři dostanou od registrační agentury přidělený kód (identifikující národní jmenný podprostor a obsahující též kód typu identifikátoru) a je plně v jejich gesci generování jedinečných řetězců NBN.

Instituce, která se chce stát registrátorem systému URN:NBN:NL, musí digitální objekty zpřístupňovat uživatelům online a současně je dodávat do dlouhodobého digitálního úložiště (tzv. „Long Term Preservation Depot“), který je uznán registrační agenturou a kde budou trvale uchovávány. V seznamu schválených úložišť⁵¹ jsou momentálně uvedena dvě dlouhodobá úložiště: e-depot⁵² (úložiště KB) a EASY.⁵³ Úložiště e-depot slouží k ukládání publikací, úložiště EASY pro výzkumná data. Odpovědnost za splnění podmínek leží na registrátorech, kteří například musí mít smlouvu s provozovatelem digitálního úložiště uzavřenu dříve, než začnou s registrací, a musí dopředu předložit přehled o tom, jaké dokumenty a kdy budou registrovat.

Institucionální repozitář (zpřístupňovací systém registrátora) musí mít zaveden protokol OAI-PMH a užívat metadatové formáty DIDL nebo MODS, koncovým uživatelům poskytovat odkazy obsahující URN:NBN (ve tvaru adresa resolveru a URN:NBN), dodržovat pravidla pro systém URN:NBN:NL a spolupracovat na kontrole kvality vykonávané registrační agenturou. Mezi pravidla systému patří především dodržování právní odpovědnosti, tedy registrovat jen takové dokumenty, k nimž má registrátor právo k zveřejnění nebo alespoň uchování. Každý digitální objekt s URN:NBN musí být současně zaslán do schváleného dlouhodobého digitálního úložiště. Stránka, na kterou následně bude resolver DANS odkazovat, by měla obsahovat podrobný popis dokumentu (např. v případě článku by zde měly být i dostupné údaje o periodiku, ve kterém vyšel). Samotný resolver DANS však uchovává jen nezbytná metadata o přidělení, o formátu dokumentu a o právní odpovědnosti za dokument. Zbylá popisná metadata jsou umístěna až v zobrazovacím systému registrátora, který je též odpovědný za jejich aktuálnost. Důvodem k přidělení nového identifikátoru je podstatná změna obsahu identifikovaného dokumentu. Registrátor si URN:NBN sám generuje (případně za použití poskytnutého nástroje) a vkládá je do metadat, která jsou pak průběžně sklízena sklízecím nástrojem identifikačního systému. Po přidělení je nutné identifikátor vložit do metadat dokumentu, dokument vystavit v institucionálním repozitáři, který musí být následně sklizen. Až poté se identifikátor stává platným, tedy v okamžiku, kdy je k němu v registru přiřazena URL adresa digitální instance. V případě, že je digitální dokument zkopírován do dalšího institucionálního repozitáře, musí být zachováno jeho URN:NBN.

Plnění podmínek je registrační agenturou kontrolováno, sleduje se dodržování oznámených plánů přidělování a dalších pravidel, především toho, aby byl digitální dokument s URN:NBN, který je zpřístupňován v institucionálním repozitáři, také uchováván

⁵⁰ Například urn:nbn:nl:kb-1382537057536.

⁵¹ <https://www.kb.nl/organisatie/onderzoek-expertise/informatie-infrastructuur-diensten-voor-bibliotheken/registration-agency-nbn/catalogus>

⁵² <https://www.kb.nl/en/organisation/research-expertise/long-term-usability-of-digital-resources>

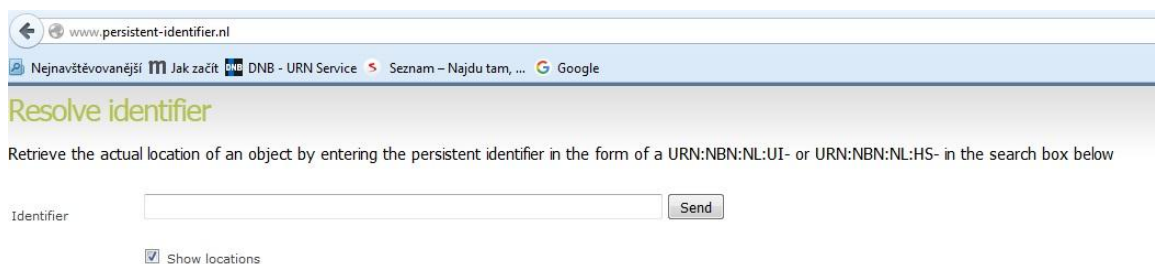
⁵³ <https://easy.dans.knaw.nl>

v dlouhodobém digitálním úložišti a evidován v registru identifikačního systému. Samotné dlouhodobé digitální úložiště musí splňovat určitá kritéria, zejména splňovat nároky dané standardem pro audit dlouhodobého digitálního úložiště „Data Seal of Approval“.⁵⁴ Dlouhodobá digitální úložiště, ať již e-Depot nebo systém EASY, mají sjednány s registrátory různé způsoby dodávání dokumentů, mezi nimi dodávání přes FTP nebo nahrávání dat přes webové rozhraní.

Technická infrastruktura

Nizozemský identifikační systém se de facto skládá ze dvou identifikačních podsystémů. Ty spojuje pouze centrální řízení jmenných podprostorů a společná pravidla (oboje zajišťované registrační agenturou). Každý podsystém má svůj vlastní resolver a registr. Jednotlivé registry obsahují pouze identifikátory URN:NBN registrátorů, kteří patří do té skupiny (viz model pro syntax), kterou daný podsystém obsluhuje. Z tohoto důvodu také jednotlivé resolvers umožňují dereferenci pouze identifikátorů URN:NBN dané skupiny.

Celonárodní resolver je dostupný na speciální doméně (<http://www.persistent-identifier.nl/>) (viz Obr. 6). Jeho registr eviduje všechny identifikátory URN:NBN registrátorů z oblastí vědeckých a výzkumných institucí (tedy takové, které jsou tvořeny podle druhého modelu pro syntax). Tento resolver spravuje organizace DANS. KB pak provozuje svůj vlastní resolver (<http://resolver.kb.nl/>) určený pro dereferenci identifikátorů digitálních publikací ukládaných do dlouhodobého digitálního úložiště e-Depot, který provozuje právě KB. [De Bruin, 2015]⁵⁵ Registrátorem těchto publikací je sama KB. KB má též svůj vlastní prezentační systém pro zpřístupňování svých dokumentů podle nastavených pravidel přístupu. [De Bruin, 2015] Resolver KB přeměrovává právě do tohoto systému. Identifikované publikace jsou však z převážné většiny volně nedostupné, neboť se jedná o komerční publikace.



Obr. 6 - Webové rozhraní resolveru DANS

Součástí identifikačního podsystému KB je generátor. Součástí druhého podsystému generátor není. Zde si registrátoři generují řetězce NBN v rámci svých přidělených podprostorů sami a ručí za jejich jedinečnost v daném podprostoru. Evidenci v celonárodním subsystému zajišťuje jeho sklízecí nástroj. Sklizením jsou evidovány nejen aktuální URL adresy digitálních instancí, ale též (při první evidenci) samotné identifikátory URN:NBN.

Německo (Švýcarsko, Rakousko)

Nejrobustnější a počtem identifikovaných dokumentů i zapojených registrátorů největší je systém URN:NBN spravovaný Německou národní knihovnou (Deutsche Nationalbibliothek, DNB, <http://www.dnb.de/>).

Systém provozovaný DNB spravuje nejen jmenný prostor URN:NBN:DE, ale i jmenné prostory pro Švýcarsko a Rakousko, které mají sice vlastní pravidla, ale infrastruktura a softwarové vybavení je stejné [Schweizerische, 2014, s. 16] (viz Obr. 7). Ze stejného

⁵⁴ <http://datasealofapproval.org/en/>. Dále DSA.

⁵⁵ Tento odkaz nebyl v době psaní článku funkční.

technického řešení plyne, že i pravidla se liší jen v detailech. Německá infrastruktura tvoří navíc základ mezinárodního clusteru, který se snaží vytvořit určitý metaresolver umožňující z jednoho rozhraní provádět dereferenci identifikátorů z více národních jmenných prostorů. V současnosti je do projektu na vytvoření clusteru kromě Německa zapojeno i Švédsko a Nizozemsko. Toto řešení je založeno na sdílení (zrcadlení) evidenčních záznamů o identifikátorech URN:NBN mezi národními registračními agenturami, nikoli na sdílení databází nebo dodržování stejných pravidel užití. Zapojené instituce si budou navzájem poskytovat pouze kopie těchto záznamů, které pak slouží pro přesměřování na aktuální URL adresy. Švýcarsko spravuje svůj jmenný prostor na vlastní infrastruktuře s využitím softwaru DNB, ale resolver je využíván společný.

S využíváním systému URN:NBN začali v DNB v roce 2001 a do roku 2005 se jednalo spíše o výzkumné práce a pilotní provoz v rámci projektu EPICUR, ale i přesto již tehdy byly identifikovány tisíce dokumentů.⁵⁶ Také v Německu byl systém zprvu zaveden v souvislosti se zaváděním povinného výtisku elektronických publikací. Bylo nutné identifikovat nově přijímané publikace. Následně byl standard URN:NBN využit pro identifikaci dalších digitálních a digitalizovaných publikací, zejména vysokoškolských prací a výsledků digitalizace odborných publikací. Systém vznikl ve spolupráci s ministerstvem vědy a výzkumu (Bundesministerium für Wirtschaft und Technologie, <http://www.bmwi.de/>). Po roce 2005 byl již systém využíván v plném rozsahu a počet identifikovaných objektů narostl na několik milionů (viz Obr. 8). Změna systému i pravidel přišla v roce 2012 v souvislosti s rozšířením typů dokumentů, které lze identifikovat.⁵⁷ V té době začal narůstat počet identifikovaných dokumentů, stejně jako přibýlo digitalizačních projektů a navýšilo se množství typů dokumentů, které byly reálně využity. To vše vedlo k úpravám a zjednodušení pravidel a optimalizaci celého systému, protože jen v německé části přesáhl v roce 2015 počet identifikovaných dokumentů výrazně mez 15 milionů a počet registrátorů převýšil 350. Švýcarská část operuje s desítkami dalších registrátorů.



Obr. 7 - Webové rozhraní URN:NBN resolveru provozovaného DNB

Účastníci

Registrační agenturou německého systému je DNB, která přiděluje kódy registrátorům a provozuje technickou infrastrukturu identifikačního systému. Registrátory se mohou stát veřejné nebo vědecké instituce nebo větší vydavatelé sídlící v Německu. Zapojení jednotlivých univerzitních ústavů nebo projektů není možné, tyto se musejí obracet na svou mateřskou

⁵⁶ <http://www.dnb.de/EN/Wir/Projekte/Abgeschlossen/epicur.html>

⁵⁷ http://www.dnb.de/EN/Wir/Sammelauftrag/sammelauftrag_node.html

instituci, která se registrátorem stát může. Dále platí, že jakákoli instituce, která je provozovatelem dlouhodobého digitálního úložiště, se musí stát také registrátorem německého identifikačního systému a přidělovat identifikátory URN:NBN dokumentům, které přijímá k dlouhodobému uchovávání (pokud tyto již identifikátor URN:NBN nemají přiděleny z dřívější doby; tj. pokud dokument předaný do úložiště není identifikovaný pomocí URN:NBN identifikátoru, správce úložiště mu přidělí identifikátor ze svého podprostoru). [Schweizerische, 2014, s. 22]. Registrátorem online publikací, které spadají do systému povinného výtisku a jsou určeny k dlouhodobému uchovávání v DNB, je právě DNB. Ta přiděluje identifikátory publikacím během přijímání do systému. [Ackermann, 2012, s. 5]

V současnosti se systému účastní již více jak 350 registrátorů.⁵⁸ Mezi registrátory jsou zejména knihovny, vydavatelé, univerzity a výzkumné ústavy.

Objekt identifikace

Německo

V německém systému platí pro objekt identifikace následující podmínky: musí tvořit samostatnou a neměnnou obsahovou jednotku, být trvale uchovávan (v případě, že dojde k jeho formátové migraci, výsledný dokument dědí původní identifikátor URN:NBN) a být dostupný (na úrovni granularity, v níž je identifikován) online. [Ackermann, 2012, s. 11-14] Obecné omezení typu dokumentu, kterému lze přidělit identifikátor URN:NBN v německém jmenném prostoru není stanoveno. Nicméně jsou vyjmenovány základní doporučené typy dokumentů a možné úrovně granularity, s poznámkou, že výčet se může rozšiřovat do budoucna a některé případy je nutno posuzovat individuálně s ohledem na výše uvedené podmínky. Aktuální výčet zahrnuje následující typy. Digitální monografie (jako příklad jsou uváděny diplomové práce nebo elektronické knihy) lze identifikovat na úrovni titulu. Periodika lze identifikovat na úrovni čísla a článku. Titul periodika může být identifikován pouze v případě, že již není vydáván a registrátor má k dispozici všechna jeho čísla. Pro identifikaci titulu dosud vycházejících periodik je doporučeno užít identifikátor ISSN nebo německý ZDB.⁵⁹ Digitalizáty tištěných dokumentů se doporučuje identifikovat na úrovni jednotlivého díla (ve smyslu knihy, čísla periodika apod.). V případě, že digitalizáty existují v různých verzích (formátech), identifikátor URN:NBN je tentýž pro všechny. Webové stránky mohou být identifikovány pouze jako objekt uchovávaný ve webarchivu (tedy obsahově již neměnný objekt). Audiovizuálním dokumentům lze také přidělit identifikátor URN:NBN. Identifikace výzkumných dat není v současnosti v německém systému řešena.

Švýcarsko

Švýcarská implementace je ve větší míře zaměřena na vysokoškolské práce (disertační a habilitační) a na archivaci webu, kdy identifikované jsou informační balíčky obsahující jednotlivé webové stránky v určitém časovém úseku (tedy archivované webové stránky). [Schweizerische, 2014] Obecné vymezení definuje, že stránky s doménou .com, .org, .info, .tv atd. obdrží URN:NBN, pokud je hlavní sídlo dané instituce ve Švýcarsku a stránky jsou vybrány některou z kantonálních knihoven nebo jinou oprávněnou institucí a dlouhodobě archivovány v národní knihovně. I stránky s doménou jiné země obdrží URN:NBN, pokud jsou dílem švýcarského autora či autorky nebo švýcarského kolektivu nebo se obsahově vztahují ke Švýcarsku a jsou vybrány některou z kantonálních knihoven nebo jinou oprávněnou institucí a dlouhodobě archivovány v národní knihovně. [Schweizerische, 2014, s. 18-20]

Kromě výše zmíněných však mohou být identifikovány i elektronické publikace, digitalizované plakáty, nebo digitalizované zvukové dokumenty. Švýcarská pravidla předepisují, že vše, co má být archivováno v úložišti Národní knihovny Švýcarska, musí být označeno pomocí identifikátoru URN:NBN, i kdyby daný dokument již měl přidělen jiný typ identifikátoru. [Schweizerische, 2014, s. 7]

⁵⁸ <http://nbn-resolving.org/institutions>

⁵⁹ Zeitschriftendatenbank (jednotný katalog periodik).

Model pro syntax

V německém systému existují dva modely pro syntax: [Ackermann, 2012]

1.	urn:nbn:de:<kód sítě knihoven>:<kód knihovny>-<řetězec NBN>
2.	urn:nbn:de:<kód registrátora>-<řetězec NBN>

První model je určen pro registrátory, kterými jsou knihovny. Příkladem je urn:nbn:de:gbv:089-3321752945. [Ackermann, 2012, s. 8-9]

Druhý model pro všechny ostatní typy registrátorů (vydavatele a různé instituce a též knihovny, které nejsou zařazeny do knihovní sítě). Kód registrátora je zde obvykle tvořen čtyřmístným číslem. Příkladem je urn:nbn:de:0292-97839421307388. Oba modely jsou obdobně využívány i pro švýcarské a rakouské prostředí. [Schweizerische, 2014, s. 9]

Řetězec NBN je ve všech případech vytvářen podle následujícího modelu:

<identifikátor dokumentu><kontrolní součet>

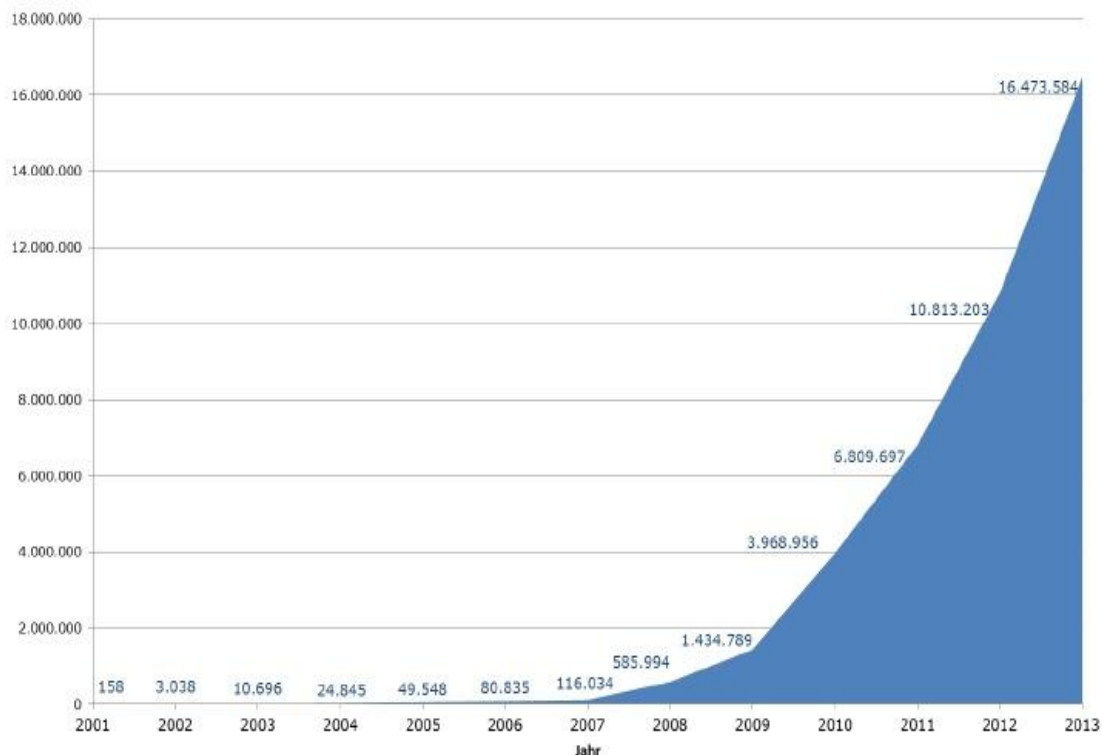
Kontrolní součet je tvořen jednou číslicí (0-9) a slouží pro detekci náhodné chyby při přepisu identifikátoru.⁶⁰ Výpočet kontrolního součtu je v tomto případě jednoduše implementovatelný a není zbytečně výpočetně náročný.

Pro prefix identifikátoru URN:NBN platí omezení, že musí být zapsán malými písmeny.⁶¹ Pokud není toto pravidlo dodrženo, resolver nedokáže provést dereferenci. Pro řetězec NBN toto omezení neplatí.⁶²

⁶⁰ <http://www.persistent-identifier.de/?link=316>

⁶¹ Tj. celá část identifikátoru URN:NBN od začátku až ke spojovníku, za kterým začíná řetězec NBN.

⁶² Viz například urn:nbn:de:swb:90-AAA2120045



Obr. 8 - Počty přidělených identifikátorů URN:NBN v systému DNB

Pravidla

Podmínkou pro to, aby se instituce mohla stát registrátorem německého systému je, aby identifikované dokumenty byly uchovávány v dlouhodobém digitálním úložišti, které prošlo certifikací minimálně podle standardu DSA, případně musí mít tato instituce zdokumentovaný záměr je v takto certifikovaném úložišti uchovávat. [Ackermann, s. 5]⁶³

Registrační agentura je zejména vázána pravidly, obsaženými v německém standardu vymezujícím kritéria pro hodnocení důvěryhodnosti systémů pro perzistentní identifikátory. [Nestor, 2009] Tato kritéria se týkají především požadavků na technické, organizační, finanční a bezpečnostní zajištění provozovatele technické infrastruktury (zejména provozu resolveru).

Registrátor si musí během kalendářního roku registrovat alespoň jeden dokument, jinak přijde o práva ke správě svého jmenného prostoru a musí o ně znovu žádat. [Ackermann, 2012, s. 6] Pokud do dlouhodobého digitálního úložiště, provozovaného registrátorem, přijde dokument, který nemá identifikátor URN:NBN, pak má tento provozovatel povinnost označit dokument identifikátorem URN:NBN ze svého jmenného prostoru. Tento požadavek je zdůvodněn tím, že každý správce dlouhodobého úložiště musí mít možnost identifikovat dodaný digitální dokument v případě, že dokument dosud nemá perzistentní identifikátor, což je nutné k naplnění potřeb dlouhodobé správy.

Registrátoři musejí do registru identifikačního systému odeslat metadata pouze v omezeném rozsahu (formát dokumentu, čas jeho registrace, vlastník, kontrolní součet). Popisná metadata musí udržovat registrátor ve svém katalogizačním nebo jiném systému.

⁶³ Pravidla pro úložiště jsou specifikována na adrese: <http://nbn-resolving.de/urn:nbn:de:0008-20080710227>

DNB přidělí identifikátor každému dokumentu, který jí je v rámci povinného výtisku nebo jinak předán a splňuje kritéria pro přidělení. Pokud vydavatel splní technické a organizační podmínky a projeví zájem, je mu přidělen vlastní jmenný prostor a může přidělovat identifikátory sám. Platí zásada, že identifikátor dokumentu zůstává, i když je kopírován do dalších systémů. Formát dokumentu a bibliografický popis nemá k identifikátoru přímý vztah, to znamená, že změna formátu nevytváří potřebu přidělení nového identifikátoru. Naopak při jakékoliv změně obsahu dokumentu je znovupřidělení nezbytné. Naopak pro bibliografický popis platí zásada, že identifikován je sám dokument, nikoli jeho popis. Všichni registrátoři musí zaručit trvalou přístupnost registrovaných dokumentů nebo ohlásit DNB, že dokument již není dostupný. Identifikátor URN:NBN musí být zároveň viditelně zapsán přímo v dokumentu nebo v údajích, které ho popisují.

K jednomu identifikátoru URN:NBN lze přiřadit více URL adres, z nichž jedna může být primární (výchozí přesměrovávání je pak na tuto primární adresu). V německém systému neplatí pravidlo, že registrátor musí být vlastníkem dokumentu. Vlastník dokumentu může požádat kteréhokoliv registrátora, aby za něj jeho dokument zaregistroval.

Technická infrastruktura

Národní resolver dostupný na adrese <http://nbn-resolving.de/> a společný i pro švýcarský a rakouský prostor na <http://nbn-resolving.org> provozuje DNB. Resolver zajišťuje dereferenci nejen německých, ale také švýcarských a rakouských identifikátorů URN:NBN a také identifikátorů DOI. Výsledkem dereference je přesměrování na aktuální URL adresu.⁶⁴ Resolver dokonce umožňuje i reverzní dereferenci – na základě URL adresy digitální instance vyhledat identifikátor URN:NBN. V tomto případě se vygeneruje stránka, na které je napsáno, k jakému URN:NBN je daná URL adresa (v registru systému) připojena.⁶⁵

Pro dereferenci je však nutné, aby písmena v identifikátoru URN:NBN v části před řetězcem NBN byla zapsána jako malá písmena. Pro část tvořenou vlastním řetězcem NBN toto omezení neplatí. Chování služby resolveru, případně formát dat získaných dereferencí, je možné dále ovlivnit volitelnými parametry.

Součástí německého identifikačního systému není obecně přístupný generátor. Každý registrátor je zodpovědný za jedinečnost svého řetězce NBN, který si vygeneroval v rámci přiděleného podprostoru. Ke kontrole správnosti vygenerovaného identifikátoru může posloužit zmíněný kontrolní součet, který tvoří poslední znak v řetězci NBN. Na webové stránce systému je rovněž k dispozici generátor kontrolního součtu.⁶⁶

V registru identifikačního systému lze k záznamu identifikátoru URN:NBN připojit více URL adres. V případě více URL adres je nutné jedno označit jako primární, na které bude uživatel vždy přesměrován. Mezi záznamy identifikátorů lze také vytvářet časové vazby typu předchůdce – následník (označení původního – nástupnického dokumentu) nebo hierarchické vztahy (např. vazba článku k jeho mateřskému periodiku). Dále lze nastavit příznak „frontpage“, pokud je přesměrováváno na metadata, popisující identifikovaný dokument.

Evidenční služba (registrace URN:NBN, vložení/odebrání/úpravy URL digitální instance, přiřazení URN:NBN pro novou verzi dokumentu apod.) funguje několika způsoby.⁶⁷ Preferovanou cestou je využití protokolu OAI-PMH 2.0, k čemuž je nutné využívat XML

⁶⁴ <http://www.persistent-identifier.de/download/EPICUR-Resolver-Dokumentation-v11.pdf>

⁶⁵ Například pro URL adresu <http://digbib.ubka.uni-karlsruhe.de/volltexte/212004> se zobrazí oznámení: „The URL <http://digbib.ubka.uni-karlsruhe.de/volltexte/212004> has following URN urn:nbn:de:swb:90-AAA2120045“.

⁶⁶ <http://nbn-resolving.de/nbnpruefziffer.php>

⁶⁷ <http://www.persistent-identifier.de/?link=220>

formát xepicur.⁶⁸ Kromě této možnosti lze registraci a další operace provádět pomocí webového rozhraní⁶⁹ a také pomocí emailové komunikace na adresu Urn-transaction@nbn-resolving.org. Registrace URN:NBN musí proběhnout před nebo při publikování dokumentu, následné přidělení není možné. Poté, co je v systému registrován identifikátor s metadaty, má registrátor povinnost do 24 hodin dodat k záznamu jedno či více URL.

Pro německý identifikační systém vznikl v rámci projektu EPICUR také plugin pro webový prohlížeč.⁷⁰ Plugin funguje jako tzv. protocol handler pro protokol URN:NBN. Pokud v prohlížeči s tímto nainstalovaným pluginem napíšeme do adresního řádku například `urn:nbn:de:gbv:089-3321752945`, prohlížeč zobrazí "<http://nbn-resolving.de/urn:nbn:de:gbv:089-3321752945>". To umožňuje do webových stránek vkládat odkazy typu `Nitric Oxide in the Olfactory Epithelium` a prohlížeč s nainstalovaným pluginem si s tím dokáže poradit. Nicméně plugin není již udržován a pro současné prohlížeče je nepoužitelný.

Součástí německého systému je i administrativní komponenta, která umožňuje přihlášení se do systému jednotlivými registrátory.⁷¹ DNB současně provozuje i mezinárodní resolver, a to na adrese <https://nbn-resolving.org/>. Ten je spojen s vlastním registrem, ve kterém je zrcadlen registr národního identifikačního systému, a tak tento mezinárodní resolver dokáže provádět rovněž dereferenci německých (a švýcarských a rakouských) identifikátorů URN:NBN. Jedním z výstupů dereference je i zobrazení (v registru uložených) všech URL adres spojených s daným identifikátorem URN:NBN nebo XML souboru obsahujícím zejména administrativní údaje o identifikátoru URN:NBN (např. datum vytvoření nebo změny záznamu).⁷² DNB též připravuje možnost provádět dereferenci identifikátorů URN:NBN několika dalších zemí. Národní knihovna České republiky v současnosti provádí testování možnosti propojení systému ČIDLO s tímto mezinárodním resolverem. Dereference je pro zahraniční resolvers prováděna obdobně jako u výše popsaného resolveru, tedy pomocí zrcadlení databází jednotlivých systémů.

Na webovém rozhraní mezinárodního resolveru jsou (na rozdíl od rozhraní národního resolveru) obsaženy základní informace o německém identifikačním systému, včetně seznamu všech německých registrátorů⁷³ a jim přidělených jmenných podprostorů.⁷⁴ Tyto seznamy jsou žel vedeny odděleně, takže není možné zjistit, jaký kód registrátora byl přiřazen té které instituci.

⁶⁸ Příklady: http://www.persistent-identifier.de/download/OAI_URN_Verwaltung_Beispiele.zip

⁶⁹ <https://ssl.nbn-resolving.de/frontend/>

⁷⁰ <http://www.persistent-identifier.de/?link=550>

⁷¹ Pro přihlášení je určen tento odkaz: <https://admin.nbn-resolving.org/>

⁷² Viz například: <http://nbn-resolving.org/resolver?identifier=urn%3Anbn%3Ade%3Agbv%3A089-3321752945&verb=full&xml=on>

⁷³ <http://nbn-resolving.org/institutions>

⁷⁴ <http://nbn-resolving.org/namespaces>

Systém ČIDLO (URN:NBN systém v České republice)

Obsahem příspěvku je především prozkoumat jednotlivé implementace URN:NBN systémů v Evropě a srovnat nástroje na jejich podporu. Není naším záměrem podrobně popisovat tento systém v České republice, protože tak již bylo učiněno jinde [Cubr, 2013; Vašek, 2014], ale pro možnost srovnání s dalšími popisovanými systémy zde budou uvedeny alespoň hlavní náležitosti české implementace, která se nazývá ČIDLO (Český systém pro Identifikaci a Lokalizaci dokumentů digitálního kulturního dědictví). Systém byl vyvinut v letech 2011-2012 jako součást interního výzkumu Národní knihovny ČR (NK ČR, <http://www.nkp.cz/>), od roku 2012 je pak nasazen v ostrém provozu a do konce roku 2015 byl identifikován více než 1 milion dokumentů. Od doby prvního nasazení je software průběžně doplňován a jsou rozšiřovány možnosti jeho funkcionalit. Nástroj je vydaný pod otevřenou licencí GNU GPL v3. Zdrojové kódy, instalační balíky, dokumentace apod. jsou veřejně dostupné na webu projektu.⁷⁵

Účastníci

Centrální agenturou tohoto systému je NK ČR, která prostřednictvím své centrální aplikace, resolveru, přiděluje identifikátory institucím (registrátorům). Ti zasílají centrální aplikaci požadované bibliografické a technické údaje a zpětně obdrží identifikátor, který zapíšou do metadat. Kromě on-line přidělování existuje možnost rezervace řady identifikátorů, které se stanou platnými po zaslání požadovaných údajů do centrální aplikace. Tento postup je plánován zejména pro případ možného výpadku infrastruktury centrální aplikace. Na straně registrátora není třeba instalace žádného softwaru, jen je třeba upravit jeho produkční prostředí tak, aby mohlo automatizovaně komunikovat s centrální aplikací. Registrátorem jsou všechny knihovny v ČR a vydavatelé elektronických knih a on-line periodik, kteří projeví zájem o zapojení do systému (povinně se to týká všech, kdo využívají Standardy Národní digitální knihovny (NDK)). U dalších institucí záleží na povaze identifikovaných objektů. Do systému je v současnosti zapojeno přibližně 100 účastníků, zejména knihovny. Kromě role registrátora může účastník získat i roli archivátora, tedy toho, kdo je odpovědný za dlouhodobé uložení identifikovaných dokumentů.

Objekt identifikace

Pro identifikované dokumenty platí ve službě ČIDLO podobná pravidla jako v německé implementaci standardu URN:NBN. Objekt identifikace (intelektuální entita) musí tvořit samostatnou a neměnnou obsahovou jednotku, která je součástí digitálního kulturního dědictví, být trvale uchovávan (v případě, že dojde k jeho formátové migraci, výsledný dokument dědí původní identifikátor URN:NBN) a být dostupný online (nebo alespoň musí být přijaté rozhodnutí o zpřístupnění objektu v blízké budoucnosti). Systém je možné využít pro digitalizované knihovní sbírky (zejména monografie, periodika a kartografické dokumenty), kvalifikační vysokoškolské práce, elektronické publikace a články publikované on-line, pokud splňují podmínku neměnnosti obsahu, a další typy objektů (např. články, kapitoly knih, ročníky periodik). Identifikovat lze více úrovní v rámci jednoho celku, např. lze přidělit identifikátor číslu periodika a následně (jiný) identifikátor článku v něm obsaženém. Volba granularity závisí na rozhodnutí registrátora, povinná je identifikace intelektuální entity, která je samostatným celkem (tedy nutná je identifikace čísla periodika, volitelná pak jeho jednotlivých článků nebo nadřazeného ročníku). V případě, že digitalizáty existují v různých verzích (formátech), identifikátor URN:NBN je tentýž pro všechny, pokud se jedná o stejného registrátora.

Model pro syntax

Syntax identifikátoru je v systému ČIDLO následující: `urn:nbn:cz:<kód registrátora>-<řetězec NBN>`, přičemž kód registrátora zpravidla odpovídá sigle knihovny, jiní registrátoři mají specifické kódy zjistitelné ze seznamu registrátorů na webu služby ČIDLO. Řetězec NBN je vždy šestimístný alfanumerický řetězec. Identifikátor musí být zapisovaný malými

⁷⁵ <https://github.com/NLCR/CZIDLO>

písmeny, ale vyhledávání v databázi a vlastní resolvování je funkční i tehdy, pokud požadavek obsahuje velká písmena.

Pravidla

Mezi základní pravidla patří podobně jako v jiných systémech odpovědnost registrátora za správnost přidělování identifikátoru, tedy za volbu správného druhu dokumentu, za odeslání odpovídajících metadat do centrální aplikace a za zapsání identifikátoru do metadat archivní i uživatelské kopie. Dále je třeba deklarovat dlouhodobé uložení identifikovaného dokumentu (splnění této podmínky není třeba dokládat), zajistit trvalou údržbu online přístupu k identifikovaným dokumentům a umožnit využívání identifikátorů URN:NBN k přesměrování na aktuální lokaci dokumentu na webu. Identifikovaný dokument je možné následně měnit jen v souladu s metodikou pro udržování životního cyklu identifikátoru, je možné doplňování chybějících popisných metadat, nesmí docházet k signifikantním změnám dokumentů. Pokud k takové změně dojde, je třeba provést v centrální aplikaci deaktivaci identifikátoru a přidělení nového, přičemž je doporučeno využít možnosti vytvoření vazby předchůdce-následník.

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. Po vystavení dokumentu v digitální knihovně nebo po publikaci na webu registrátor připojí k identifikátoru v databázi systému ČIDLO jednu nebo více URL adres (digitální instance), na kterých je objekt umístěn. Aktuálnost adres je průběžně kontrolována a případně opravována automatickým procesem sklizením metadat přes OAI-PMH pomocí komponenty OAI Adapter. V případě, že registrátor opomene zaregistrovat URL adresu, je doplněna při pravidelné automatické sklizni registrované digitální knihovny. Registrátor je nadále oprávněn spravovat svůj jmenný prostor, tedy může v rámci pravidel řízení životního cyklu dokumentu provádět úpravy dokumentů a digitálních instancí. K některým operacím je třeba spolupráce správce centrální aplikace. Dereference identifikovaného dokumentu se provádí pomocí schématu <adresa resolveru>/<identifikátor URN:NBN>. Všechny identické kopie digitálního dokumentu obdrží tentýž identifikátor URN:NBN. Dereference jednoho URN:NBN tudíž může zahrnovat přesměrování na několik odlišných URL adres, pořadí určuje nastavení systému. V případě, že dokument není dostupný (což by se běžně stávat nemělo), je možné v databázi vyhledat identifikátor a k němu připojený bibliografický záznam, což může zpětně sloužit k ověření existence dokumentu.

Technická infrastruktura

Systém ČIDLO je založen na jediném celonárodním resolveru s centrální aplikací a databází. Součástí systému není obecně přístupný generátor. Každý registrátor spravuje svůj jmenný prostor, jedinečnost identifikátoru zajišťuje centrální aplikace. V registru identifikačního systému lze k záznamu identifikátoru URN:NBN připojit více URL adres. Mezi záznamy identifikátorů lze také vytvářet časové vazby typu předchůdce – následník (označení původního – nástupnického dokumentu) nebo hierarchické vztahy (např. vazba článku k jeho mateřskému periodiku). Samotný resolver je dostupný na adrese <https://resolver.nkp.cz>. Jeho webové rozhraní je poměrně rozsáhlé, nabízí sekci pravidel, seznam registrátorů a vlastní vyhledávání. Po přihlášení umožňuje další funkce, především z hlediska správy práv a uživatelských účtů, dále poskytuje možnost generování identifikátoru přes formulář webového rozhraní a provádění dalších operací. Resolver umí vykonávat pouze standardní dereferenci na URL digitální instance a to jak pomocí přiřazených URL adres, tak přímo pomocí výše popsání schématu. Automatizovaný i webový přístup k uživatelskému účtu probíhá pomocí loginu a unikátního hesla.

Evidenční služba využívá faktu, že generátor je spojen s registrem resolveru, registrátoři proto nemají další povinnost informovat o přidělených identifikátorech. Při první registraci může registrátor zadat URL digitální instance automaticky pomocí API nebo dodat přes

webové rozhraní. Případně může vyčkat, než bude jeho digitální knihovna sklizena pomocí nástroje OAI Adapter.

Softwarové řešení je založeno na opensource technologiích (Java, databáze PostgreSQL, Spring security, GWT aj.). Systém využívá standard OAI-PMH jak pro aktivní sklizení zaregistrovaných digitálních knihoven, tak umožňuje své vlastní sklizení dalšími systémy. Automatizovaná komunikace s dalšími nástroji probíhá pomocí API.

Srovnávací analýza

Účastníci a objekt identifikace

	Finsko	Itálie	Nizozemsko	Německo	Česko
Registrační agentura	Kansalliskirjasto (FNL)	BNCF	Registration Agency NBN (KB)	DNB	NK ČR
Počet registrátorů	50+	15	50+	350+	100+
Počet URN:NBN (v milionech)	0,3	0,015	25	15	1
Omezení registrace	NE	ANO	ANO	ANO	ANO
Převažující typ registrátora	univerzity	univerzity	vysoké školy	vysoké školy	knihovny
Seznam registrátorů na webu	NE	ANO	ANO	ANO	ANO

Tab 2 Účastníci a objekt identifikace

Ve všech případech je registrační agenturou národní knihovna dané země, jak lze také očekávat, neboť to je podmínka standardu pro URN:NBN. Pouze v jediném případě však existuje uvnitř knihovny zvláštní organizační útvar pro přidělování URN:NBN (Registration Agency NBN v KB). V Itálii, kde funguje více národních knihoven současně, zastává roli národní registrační agentury URN:NBN pouze jedna z nich (BNCF). Všechny národní knihovny zkoumaných systémů jsou samy rovněž v roli jejich registrátorů (v případě Nizozemska a Česka také největších registrátorů).

Z kvantitativního hlediska za ostatními výrazně zaostává Itálie, což lze částečně chápat vzhledem k relativně nedávnému spuštění italského systému a dále vzhledem k omezení na dokumenty ukládané do národního úložiště MD. Ve všech zemích, s výjimkou Česka, dominují v roli registrátorů vysoké školy.⁷⁶ Naopak v Česku je dosud (až na několik výjimek) dominujícím typem registrátora knihovna.

Identifikátory URN:NBN v sledované oblasti nemůže používat jakákoliv instituce. Všechny národní registrační agentury – s výjimkou Finska – mají zavedena specifická omezení na typ registrátora, který může vstoupit do systému.⁷⁷ V Německu se může stát registrátorem pouze veřejná nebo vědecká instituce, případně větší vydavatel sídlící v zemi; V Nizozemsku pouze vědecká nebo vysokoškolská instituce. Z pravidel pro přidělování URN:NBN v Itálii vyplývá, že registrátorem se může stát pouze taková instituce, která (na základě povinného nebo

⁷⁶ V případě Německa jde často specificky o knihovny vysokých škol.

⁷⁷ Pokud pomineme všeobecně platné omezení, že registrátorem musí být instituce.

dobrovolného výtisku) odevzdává digitální dokumenty do národního úložiště MD (obvykle tedy vydavatel). V Česku platí pravidlo, že registrátorem se může stát jakákoliv knihovna a dále instituce pečující o kulturní nebo vědecké digitální dědictví, pokud splňuje některá další kritéria.

Podle standardu URN:NBN musí být na webu registrační agentury dostupný seznam všech registrátorů. Tento požadavek nespĺňuje – možná překvapivě – pouze Finsko.

Nejčastějšími typy identifikovaných dokumentů jsou univerzitní publikace (Finsko, Itálie, Německo), výzkumná data (Finsko, Nizozemsko) a elektronické publikace komerčních vydavatelů (Nizozemsko, Německo). Výjimkou je Česko, kde je zatím hlavním typem digitalizát tištěného dokumentu (monografie nebo periodika). Nikde jinde, s výjimkou Německa a Finska, se identifikátory URN:NBN digitalizátům dosud nepřidělují. Jejich přidělování se plánuje v Itálii a Nizozemsku.

Model pro syntax

	Finsko	Itálie	Nizozemsko	Německo	Česko
Bez podprostoru	ANO	NE	ANO	NE	NE
Podprostor 1. stupně	ANO	ANO	ANO	ANO	ANO
Podprostor 2. stupně	NE	NE	ANO	ANO	NE

Tab 3 Model pro syntax

Ačkoliv jsme zkoumali pouze 5 oblastí, užití národní modely jsou rozmanité. Standard pro URN:NBN umožňuje užití dvou základních modelů – bez jmenného podprostoru a se jmenným podprostorem. Jmenný podprostor je možno dále členit, v tabulce jsou uvedeny dvě nalezené varianty (tj. s členěním do první, resp. druhé úrovně podprostoru).

Český a italský národní syntax pro URN:NBN je homogenní – vyskytuje se právě jeden model, a sice se jmenným podprostorem prvního stupně.⁷⁸ Podle tohoto modelu jsou v Česku a Itálii generovány všechny identifikátory URN:NBN. V Nizozemsku se vyskytují všechny výše uvedené modely, nicméně první dva se týkají pouze subsystému KB. Pro ostatní nizozemské instituce je závazný model, který využívá první podprostor k identifikaci skupiny registrátora (výzkumné instituce / vysoké školy) a druhý k identifikaci konkrétní instituce (tedy registrátora dané skupiny). Německá syntax je sice heterogenní, ale přehledná (pokud se uživatel vyzná v německé síti knihoven) – model s podprostorem 2. stupně je určen pouze pro registrátory zařazené do knihovní sítě, model s podprostorem 1. stupně pro všechny ostatní instituce.

Ve Finsku jsou užití oba základní modely syntaxe, ovšem dosti nepřehledně. Tato skutečnost je způsobena historickým vývojem (první užití identifikátoru URN:NBN v praxi) a užitou technologií. Pro všechny registrátory, kteří si přidělují identifikátory URN:NBN sami (tj. bez užití online generátoru), platí model s podprostorem 1. stupně, kdy řetězec NBN lze vytvořit z libovolných alfanumerických znaků. Potud je syntaxe přehledná a stejná jako v Česku a Itálii. Situace se však komplikuje s online generátorem, které finský systém nabízí. Pokud registrátor nechce mít svůj vlastní podprostor (a tedy si nebude generovat identifikátory URN:NBN sám), pak musí užít model syntaxe bez podprostoru, na jehož základě mu generátor přidělí identifikátor. To působí poněkud zvláště, když si uvědomíme, že vlastní národní prostor je delegován komukoliv, kdo užije generátor. Nemusí to být ani registrátor (ale např. autoři tohoto článku, kteří si zkusmo vygenerovali několik identifikátorů URN:NBN). Platným se sice identifikátor URN:NBN ve Finsku stává až po evidenci

⁷⁸ urn:nbn:<kód země>:<kód registrátora>-<řetězec NBN>

v národním identifikačním systému, nicméně toto nastavení vede k tomu, že kdokoliv bez vztahu k systému si může nechat přiřadit libovolný počet URN:NBN, a tím zabírat pozice skutečným registrátorům. Situaci dále komplikuje podvojná možnost nabízená čtyřem vybraným registrátorům, kteří si buď mohou nechat pomocí online generátoru přidělit identifikátory URN:NBN v rámci svých přednastavených podprostorů, nebo si mohou identifikátory generovat sami. Online generátor navíc vnucuje řetězci NBN speciální strukturu, obsahující významové prvky (datum).

Celkově lze říci, že italský a český model je nejjednodušší, neboť zde je syntax jednotná pro všechny registrátory a její užití je tak nejsrozumitelnější. Z hlediska škálovatelnosti při velkém počtu registrátorů různého typu má však jistě výhody i model se členěním do druhé úrovně podprostoru (Nizozemsko, Německo), který umožňuje třídění registrátorů do skupin. Skutečnost, že – až na uvedené výjimky – všechny identifikátory URN:NBN zkoumaných systémů obsahují kód registrátora, přispívá k rychlé identifikaci vlastníka dokumentu, resp. registrátora systému. Jde de facto o analogii s modelem ISBN.

Pravidla

	Finsko	Itálie	Nizozemsko	Německo	Česko
Pravidla na webu	ANO	ANO	ANO	ANO	ANO
Identifikace obsahu (ne formátů)	ANO	neřešeno	implicitně	ANO	ANO
Vazba na archivaci	NE	ANO	ANO	ANO	ANO

Tab 4 Pravidla

Jak jsme viděli, pravidla jsou stanovena různě – v různé míře podrobnosti a s užitím různých kategorizací objektu identifikace. Ve všech případech jsou pravidla zveřejněna na webu.

Nejpropracovanější pravidla nalezneme v německém a českém systému, rozvinutá pravidla má i nizozemský systém. Co se týče stanovení toho, jaký dokument může / nemůže dostat identifikátor URN:NBN, lze provést zobecnění, že všechna sledovaná pravidla de facto primárně vycházejí z typů svých registrátorů a podle nich vyvozují typy dokumentů, které tito registrátoři mohou produkovat, resp. zpřístupňovat. Na tyto předpokládané typy dokumentů pak národní systémy ukládají svá omezení a podmínky. Tomuto zobecnění odpovídá i skutečnost, že finský systém, který je nejméně restriktivní na typ registrátora, má také nejvolnější definici dokumentu, kterému lze identifikátor URN:NBN přidělit.

Pro identifikaci periodik je vždy klíčové stanovení granularity identifikace. Ve zkoumaných systémech převažuje číslo a článek periodika. Česká a německá periodika výslovně zakazují přidělení identifikátoru URN:NBN titulu periodika.⁷⁹

Z hlediska vymezení objektu identifikace je ve finských, německých a českých pravidlech podmínka, že identifikátor URN:NBN zůstává dokumentu zachován do té doby, dokud nedojde k signifikantní změně jeho obsahu. Explicitně je v českých a německých pravidlech uvedeno, že po formátové migraci zůstává identifikátor URN:NBN zachován. Z užití v Nizozemsku vyplývá, že dvě různé formátové verze téhož dokumentu (např. článek vydaný současně v HTML i PDF) mají tentýž identifikátor URN:NBN. V českých pravidlech je toto řešeno obdobně: archivní i uživatelská verze dokumentu (každá je uložena v jiném profilu formátu JP2) sdílí totéž URN:NBN. Všechny systémy (vyjma Itálie) tak de facto užívají identifikátor URN:NBN k identifikaci informačního obsahu, nikoliv datového objektu, který tento obsah reprezentuje (v termínech normy ISO 14721). Takovéto pravidlo je tak velkým

⁷⁹ S jednou výjimkou v případě Německa.

přínosem jak z hlediska dlouhodobého uchovávání (kdy formáty dokumentů se mohou měnit, ale obsah by měl být zachován, a tedy i identifikace by měla zůstat stejná, aby přetrvávala funkce perzistentního identifikátoru), tak z pohledu zpřístupňování uživatelům (kdy jsou čtenářům nabízeny různé formátové verze téhož dokumentu). Z těchto důvodů je v praxi zkoumaných systémů dereference často užitá tak, že v případě, kdy má dokument více formátových verzí, odkazuje URL (připojené k URN:NBN v registru systému) nikoliv na jednu z verzí, ale na popisnou stránku, z níž vedou odkazy na různé formátové (tak je tomu např. u článků v Nizozemsku). V Česku je popisná stránka využívána v zpřístupňovacích systémech Kramerius, v nichž z popisné (úvodní) stránky obsahující bibliografické údaje vedou odkazy na jednotlivé části dokumentu (např. naskenované stránky monografie).

V českém systému je zavedena povinnost přidělit identifikátor již při produkci, tedy na začátku životního cyklu digitálního dokumentu. Systémy, ve kterých si identifikátory URN:NBN přidělují samotní registrátoři a teprve později je dodávají do identifikačního systému, de facto pracují s podobnou premisou (přidělení na začátku životního cyklu) – dokumenty jsou vydávány s již přiděleným URN:NBN.

Ve všech národních systémech platí, že registrátoři jsou vlastníky identifikovaných dokumentů, ale v rámci těchto systémů existují jisté výjimky. V Německu může identifikátor URN:NBN přidělit provozovatel dlouhodobého úložiště (který nemusí být vlastníkem dokumentu), pokud jemu dodaný dokument identifikátor URN:NBN ještě nemá. Specifickým případem jsou národní knihovny Německa a Nizozemska, které jsou registrátory pro elektronické publikace vydavatelů, které získávají na základě povinného či dobrovolného výtisku. V tomto případě však identifikátor URN:NBN slouží spíše jen pro vnitřní potřebu správy v dané instituci – vydavatel sám jej nevyužívá.

S výjimkou Finska též všechny zkoumané systémy požadují, aby byl identifikovaný digitální dokument nejen zpřístupňován uživatelům (zpravidla v systému registrátora, např. v institucionálním repozitáři),⁸⁰ ale jeho kopie také dlouhodobě uchovávána. Stanovené podmínky na zajištění tohoto uchovávání se liší. V Nizozemsku musí být identifikovaný dokument současně dodán do dlouhodobého úložiště, a to takového, které je schváleno Národní agenturou pro NBN; v Německu do dlouhodobého úložiště, které je certifikováno minimálně podle standardu DSA. V Česku má registrátor minimálně povinnost mít alespoň záměr nechat své dokumenty v budoucnu uchovávat v dlouhodobém úložišti (toto volnější pravidlo souvisí se skutečností, že většina institucí v ČR dlouhodobé uchovávání digitálních dokumentů teprve začíná řešit). V Česku je dále zavedeno pravidlo, že pokud chce registrátor uchovávat své digitalizáty přímo v LTP úložišti NK ČR, musí je vytvářet podle standardů digitalizace NK ČR, které zahrnují také povinnost užití identifikátoru URN:NBN. V Itálii je otázka dlouhodobého uchovávání řešena z principu fungování systému samotného – identifikátor URN:NBN se přiděluje pouze těm dokumentům, které jsou dodávány k dlouhodobému uchovávání do národního úložiště MD.

Všechny systémy také vyžadují dodání metadat, která se váží k identifikátoru nebo identifikovanému dokumentu, rozsah metadat vyžadovaných jednotlivými systémy se liší, příkladem minimalistických požadavků je Finsko. Nejrozsáhlejší sadu metadat vyžaduje český systém ČIDLO, což je jeho předností, ale i slabinou, jelikož vyžaduje poměrně náročnou koordinaci při udržování životního cyklu identifikátoru. Ostatní systémy většinou pracují s premisou, že bibliografický i technický popis identifikovaných dokumentů je k dispozici v externích specializovaných systémech, do kterých se lze odkazovat. Bibliografické údaje jsou v nich omezeny na minimum, především kvůli náročnosti úkolů spojených s procesem následných aktualizací.

⁸⁰ V některých případech je toto zpřístupnění nevyhnutelně omezeno právy duševního vlastnictví, viz uvedené případy identifikace publikací vydavatelů v národních knihovnách nebo digitalizáty v Česku.

Technická infrastruktura

	Finsko	Itálie	Nizozemsko	Německo	Česko
Počet resolverů	1	1	2	2	1
Generátor	online generátor, registrátor	součást systému	součást subsystému (KB) / registrátor (DANS)	registrátor	součást systému
Evidenční služba	OAI-PMH / vystavené XML	RESTful API / plugin	OAI-PMH	OAI-PMH / webové rozhraní systému / email	OAI-PMH, RESTful API, webové rozhraní systému

Tab 5 Technická infrastruktura

Všechny zkoumané systémy provozují resolvers, Nizozemsko a Německo dokonce dva. V případě Nizozemska je identifikační systém rozdělen do dvou subsystémů sloužících odděleně jiným skupinám registrátorů, z nichž každá má svůj vlastní resolver i registr. Identifikátory z jednoho subsystému tak nelze podrobovat dereferenci v druhém. Oba německé resolvers mají tentýž registr, tedy oba umožňují provést dereferenci pro identifikátory URN:NBN Německa, Švýcarska a Rakouska. Mezinárodní resolver zatím v běžném provozu neprovádí dereferenci identifikátorů dalších zemí.

Všechny resolvers mají webové rozhraní, s výjimkou finského. Základní podobou dereference pro všechny systémy je přesměrovávání na aktuální URL adresu, na níž je dokument zpřístupňován uživateli. Ve všech případech lze také vytvořit „permanentní“ URL ve tvaru <URL resolveru><konkrétní identifikátor URN:NBN>. Německý resolver dokáže provádět i zpětnou dereferenci (z URL na URN:NBN) a dereferenci identifikátoru DOI. Finský resolver dokáže podrobit dereferenci i identifikátory URN:ISBN.

Resolvers dále nabízejí i poskytnutí přidružených metadat nebo výpis přiřazených adres URL. Pouze český systém však nabízí sadu bibliografických metadat popisujících identifikovaný dokument. Český a německý systém dále také umožňují vytvářet vazby mezi jednotlivými identifikátory URN:NBN.

Překvapivé je, že ne ve všech zemích je resolver schopen provést dereferenci bez ohledu na velká / malá písmena.

Centralizované systémy jsou pouze v Česku a Itálii (a subsystém KB v Nizozemsku), kde je přidělování zajišťováno vlastním systémem národní registrační agentury. Generátory jsou zde propojeny s registrem, což znamená, že identifikační systém o přiděleném URN:NBN okamžitě ví. V Česku i Itálii je přidělování realizováno primárně komunikací mezi systémy registrátora a registrační agentury prostřednictvím RESTful API.

V ostatních zemích je užito distribuované řešení, tedy generování identifikátorů URN:NBN je v odpovědnosti daného registrátora; přesněji řečeno jde jen o generování řetězce NBN - prefix je registrátorům přidělen registrační agenturou. V okamžiku přidělení tak identifikační systém o existenci nového identifikátoru neví. Zvláštním případem je finský online generátor, který je sice provozován přímo národní agenturou, nicméně není propojen s registrem systému, což je poměrně zvláštní situace.

Z výše uvedeného je pak logické, že evidenční služba v distribuovaných systémech eviduje (na základě komunikace mezi identifikačním systémem a systémy registrátorů) adresy URL i identifikátory URN:NBN, zatímco v centralizovaných systémech jde pouze o evidenci adres URL vztahujících se k identifikátorům URN:NBN, které jsou v okamžiku evidence již obsaženy v registru systému. Pro evidenční službu se v Itálii primárně užívá RESTful API, ve všech ostatních případech OAI-PMH. Jako alternativní varianty evidence se dále vyskytují: vystavení XML na webu registrátora, které pak identifikačním systémem sklízí (Finsko); užití webového rozhraní systému (Česko, Německo); zaslání e-mailem (Německo). V Česku je možné pro evidenci adres URL užít i RESTful API, v Itálii vyvinuli plugin pro vybrané redakční systémy vydavatelů, který komunikuje s centrálním systémem.

Nejkomplexnějším identifikačním systémem je německý, následuje český a italský. Všechny tyto tři systémy obsahují administrativní komponentu pro správu účtů ze strany registrátorů. Nejjednodušší je finský systém, kde resolver nemá ani webové rozhraní.

Z hlediska technického řešení byly zjištěny nedostatky ve většině systémů. Finský resolver funguje korektně, až na drobný nedostatek – nerozlišuje nevalidní a neexistující URN:NBN. Není také jasné, jak se systém zachová, když registrátor odstraní vazbu URN:NBN/URL ze svých metadat. V Itálii lze vidět drobné omezení v tom, že každému účtu je přiřazen právě jeden registrátor. Problematická se zdá být též nemožnost ovlivnit chování služby parametry (případně podobou URL) a vynutit třeba přesměrování přímo na URL adresu digitální instance. Nizozemský resolver provozovaný DANS trpí typickým nešvarem – z odpovědi nelze jednoduše strojově zjistit typ chyby a je nutné zpracovat (parsovat) vrácenou webovou stránku. Reakce německého (národního) resolveru na chybové stavy je ze všech porovnávaných identifikačních systémů nejlepší z pohledu strojového zpracování. Celkově tento resolver také poskytuje nejvíce funkcí. Nevýhodou českého systému se jeví jeho komplexita a tedy i značná náročnost na správu, některé z postupů nejsou zcela intuitivní a vyžadují podrobnou znalost fungování, aby mohl být plně využit.

Specifika systémů

Zkoumané systémy vykazují některá specifika, která je na tomto místě dobré shrnout.

Ve Finsku je jako v jediné ze zkoumaných oblastí současně zaveden systém URN:ISBN, přičemž platí, že když lze dokumentu přidělit URN:ISBN, URN:NBN se nepřiděluje.

Italský systém je jako jediný zcela spjat s institutem povinného (resp. dobrovolného) výtisku a dlouhodobým uchováváním v úložišti národní knihovny. Představuje vlastně subsystém italského národního úložiště MD. Italský systém je také spojen s institutem národní bibliografie (co má identifikátor URN:NBN, je součástí národní bibliografie), čímž je jediným přísným pokračovatelem původní tradice identifikátorů NBN jakožto identifikátorů výhradně národní bibliografie. Po přidělení URN:NBN provádí úložiště MD (nikoliv identifikační systém) stažení samotného dokumentu a jeho metadat a jejich uložení pro účely dlouhodobého uchovávání. V případě, že dokument již nebude v systému registrátora dostupný pro uživatele, tak vždy ještě existuje jeho kopie uložená v úložišti MD (přístupná pro oprávněné uživatele).

Nizozemským specifikem je provozování dvou de facto oddělených identifikačních subsystémů – dvou různých resolverů i registrů. Společné mají pouze řízení národní agenturou v oblasti přidělování jmenných podprostorů a stanovování pravidel. Dalším specifikem je kontrola kvality, vykonávaná rovněž národní agenturou.

Německo je jedinečné zejména v tom, že se ujal role vybudovat mezinárodní metaresolver a také počtem identifikovaných dokumentů. Jako jediný ze zkoumaných systémů umožňuje německý reverzní dereferenci a dereferenci identifikátorů DOI, a to v podobě přesměrování na aktuální URL adresu, kde je dokument označený tímto identifikátorem zpřístupňován (což

je v souladu s pravidly Systému DOI vždy úvodní stránka, popisující identifikovaný dokument).

Český systém je specifický v ukládání nejrozsáhlejších metadat o dokumentu. V případě, že dokument již neexistuje, je český resolver nástrojem, který může prokázat minulou existenci citovaného dokumentu.

Závěr

Historicky první implementace standardu URN:NBN do konkrétního národního identifikačního systému vznikly ve Finsku a Německu. Jejich počátky se překrývají s vydáním standardu pro URN:NBN. Od roku 2007 je budován identifikační systém v Itálii a Nizozemsku, přičemž v Itálii šlo až do roku 2014 pouze o konceptualizaci systému, která prošla několika etapami. Český identifikační systém ČIDLO je tak spolu s italským nejmladším (podzim 2012). Je přitom pozoruhodné, že zatímco italský systém je relativně malý, v oblasti teorie jsou to právě Italové, kteří patří mezi nejvíce publikující odborníky o problematice trvalé identifikace. Podobně zajímavá je skutečnost, že Finsko, ačkoliv standard URN:NBN vzešel právě odsud, má svůj systém na nejméně propracované úrovni ze všech sledovaných zemí.

Všechny zkoumané identifikační systémy – vyjma českého - byly zprvu výhradně zaměřeny na online publikace nebo výzkumná data z oblasti vysokého školství, výzkumu a vývoje, v případě Nizozemska (KB) navíc na identifikaci pravděpodobně největší světové sbírky elektronické produkce (knihy a články periodik) pocházející z odborných vydavatelství (ať jde již o e-borny nebo elektronické verze tištěných knih a periodik) získaných na základě dobrovolného elektronického výtisku. To se adekvátně odrazilo i v největším počtu přidělených identifikátorů URN:NBN ze všech sledovaných zemí. Připomeňme, že právě v KB byly provedeny první pokusy o vytvoření dlouhodobého digitálního úložiště, které právě souvisely se zmíněnou koncentrací elektronické vydavatelské produkce v nizozemské národní knihovně. Národní knihovny sledovaných zemí tedy od počátku milénia začaly plnit funkci, kterou v minulosti nikdy nevykonávaly – zajišťovat trvalou identifikaci specifického obsahu (vysokoškolské závěrečné práce, výzkumná data apod.), který předtím – tj. ve své tištěné reprezentaci – nebyl nikdy předtím předmětem národních bibliografií, ani identifikačních systémů ISBN nebo ISSN. Kromě prosté identifikace začaly národní knihovny též zajišťovat novou funkci identifikátoru pro digitální dokumenty, a sice přesměrovávání na aktuální lokaci dokumentu na internetu (založenou na udržování aktuálních informací o těchto lokacích). Využily k tomu standard URN:NBN, čímž vznikla obdoba identifikátorů Handle a DOI. Požadavek na tyto služby vzešel ve většině případů nejprve z oblasti mimo knihovnickou komunitu (zejména z oblasti výzkumu a vysokého školství). Knihovnická komunita v zavedení identifikačních systémů založených na URN:NBN jednak spatřovala budoucnost pro své jádrové čtenáře (tedy čtenáře jí vlastněných dokumentů v digitální podobě, které se v řadě případů začaly identifikovat až mnohem později po spuštění systémů), jednak tím předjímal očekávané zavádění povinného elektronického výtisku.

Postupem času se záběr typů identifikovaných dokumentů rozšířil na další typy dokumentů (např. v Německu na digitalizáty a povinný elektronický výtisk), nicméně dosud platí, že převažujícím objektem identifikace ve zkoumaných systémech (vyjma Česka) jsou e-borny pocházející z vysokých škol a vědeckých institucí.

Vývoj zkoumaných zahraničních systémů založených na URN:NBN také ukazuje, jak se hranice mezi univerzem knihovnických dokumentů a dokumentů z jiných oblastí kulturního a vědeckého dědictví v digitálním světě stírají.

Centralizované řešení, užití v Česku a Itálii, má nepochybně více kladů než záporů. Poskytuje větší záruku (autoritou jedné registrační agentury i vhodným technologickým řešením), že nikdy nebude přidělen duplicitní nebo chybný identifikátor URN:NBN a že je identifikátor okamžitě po vygenerování již obsažen v registru systému. Registrátoři nemusejí pracovat s dalším softwarem ani provádět další rozsáhlé investice do infrastruktury, odpadá jim i povinnost aktivně zajišťovat následné registrace v centrální databázi registru, ale i tak musí jejich systém disponovat nástroji pro sklizení, jako je OAI-PMH protokol. Větší nároky jsou naopak kladeny na registrační agentury, jejichž úkolem je celý systém udržovat, bdít nad udržováním aktuálních odkazů a poskytovat další podporu. Nevýhodou je naopak nutná spolupráce s dílčími repositáři, pokud nejsou identifikované dokumenty zároveň uloženy v dlouhodobém úložišti spojeném s centrální agenturou. Distribuované řešení přináší větší riziko chyb a dále má ten důsledek, že po nějaký (byť sebekratší čas) bude identifikátor URN:NBN neznám samotnému systému. Čtenář se tak může ve zpřístupňovacím systému registrátora setkat s dokumentem, u kterého bude uvedeno URN:NBN, které ale nebude moci hned podrobit dereferenci v resolveru. V distribuovaných řešeních Itálie a Česka je dereference možná okamžitě po přidělení URN:NBN. Současně je ale třeba rozumět důvodům užití distribuovaného řešení. Je jím zejména skutečnost, že centrální přidělování může klást vyšší technické nároky na komunikaci systému registrátora s centrálním systémem registrační agentury než pozdější evidence přes OAI-PMH. Výhodou distribuovaného řešení je také zrychlení celého procesu přidělování a operativnější kontrola na straně registrátorů. Naopak nevýhoda spočívá v nárocích na pracovníky registrátorů a na technické vybavení všech zapojených.

V identifikačních systémech zdaleka nejvíce převažuje užití jmenných podprostorů. To je vcelku logické vzhledem k tomu, že toto užití umožňuje jednodušší správu a přehlednost (zjištění vlastníka dokumentu ze samotné syntaxe).

Italský, německý, nizozemský i český systém jsou založeny na premise, že identifikátor je přidělován pouze dokumentu, jehož kopie je zároveň uchovávána v nějakém dlouhodobém digitálním úložišti. Tato premisa je klíčová, neboť vynakládat úsilí, které je zapotřebí pro správný chod identifikačního systému, na identifikování pomíjivých (ať již hodnotou nebo v důsledku špatného zajištění) dokumentů, se zdá být nerozumným počínáním.

Rovněž převažující důraz na trvalou identifikaci informačního obsahu, nikoliv datového objektu nebo aktuálního umístění, je cesta správným směrem a v souladu se strategií dlouhodobého uchovávání i trvalého zpřístupňování čtenářům.

Český systém ČIDLO je výjimečný v tom, že vznikl nikoliv z vnějšího požadavku vysokých škol nebo vědeckého sektoru, ale z vlastních potřeb knihoven samotných. Byl zaveden za účelem trvalé identifikace digitalizovaných dokumentů vznikajících v rámci rozsáhlého projektu Vytvoření Národní digitální knihovny a pro digitalizace vznikající v podprogramu VISK 7 i dalších projektech knihoven. Nicméně je připravováno rozšíření i na další oblasti, tak jak jsou obvyklé v dalších zkoumaných zemích. V Itálii a Nizozemsku se digitalizátům teprve chystají přidělovat identifikátory URN:NBN. To je poněkud překvapivé zejména v případě KB, která je známa rozsáhlými digitalizačními projekty.

Specifický je systém ČIDLO také tím, že je úzce propojen se standardizací digitalizace, kterou zajišťuje NK ČR pro všechny knihovny a kterou se také drtivá většina digitalizujících knihoven řídí. Tato standardizace, jejíž detailní specifikace jsou veřejně vystaveny na webu (což není běžné pro digitalizační projekty ve světě), umožňuje transparentní produkci a identifikaci a usnadňuje tak následné dlouhodobé uchovávání digitalizátů.

Seznam použité literatury

- Ackermann, Uta, Berner, Christiane, Elbert, Natalie et al. 2012. *Policy for issuing URNs in the urn:nbn:de namespace* [online]. Version 1.0. Leipzig; Frankfurt am Main: Deutsche Nationalbibliothek, 2015, Last updated: 29 November 2012 [cit. 2015-08-28]. Dostupné z: <http://nbn-resolving.org/urn:nbn:de:101-2013121001>.
- Altenhoener, Reinhard, Boulderstone, Richard, Steenbakkens, Johan et al. 2007. *CENL Task Force on Persistent Identifiers: Report 2007* [online]. Frankfurt am Main: CENL, 2007 [cit. 2010-05-01]. Dostupné z: http://web3.nlib.ee/cenl/docs/CENL_Taskforce_PI_Report_2006.pdf.
- Bellini, Emanuele, Bergamin, Giovanni, Messina, Maurizio et al. 2014. 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>.
- Bellini, Emanuele, Cirinnà, Chiara, Lunghi, Maurizio et al. 2010. Persistent identifier distributed system for digital libraries. *International Cataloging & Bibliographic Control* [online]. 2010, vol. 39, issue 2, p. 30-36 [cit. 2015-07-20]. ISSN 1011-8829. Dostupné z: <http://search.ebscohost.com/>.
- Bellini, Emanuele, Cirinnà, Chiara, Lunghi, Maurizio et al. 2012. Il progetto National Bibliography Number Italia (NBN:IT): Un identificatore persistente a supporto del deposito legale nazionale delle risorse digitali. *JLIS.it* [online]. 2012, Vol. 3, no. 1 [cit. 2015-07-20]. urn:nbn:it:unifi-3866. ISSN 2038-1026. Dostupné z: <http://leo.cineca.it/index.php/jlis/article/view/4789>.
- Bellini, Emanuele, Cirinna, Chiara, Lunghi, Maurizio. 2008. *Persistent identifiers for cultural heritage* [online]. Glasgow: DigitalPreservationEurope (DPE), 2008 [cit. 2012-05-30].
- Bellini, Emanuele. 2015. *NBN info* [online]. Message to: Iveta LODROVÁ. 2015-01-16 [cit. 2015-06-20]. Osobní komunikace.
- Cubr, Ladislav, Vašek, Zdeněk. 2013. Identifikátory digitálních dokumentů se zaměřením na systém URN:NBN v ČR. *Čtenář: měsíčník pro knihovny* [online]. 2013, roč. 65, č. 6 [cit. 2016-04-12]. ISSN 1805-4064. Dostupné z: <http://ctenar.svkkk.cz/clanky/2013-roc-65/6-2013/identifikatory-digitalnich-dokumentu-se-zamerenim-na-system-urn-nbn-v-cr-113-1484.htm>.
- De Bruin, Rob. 2015. *URN:NBN policies in the Netherlands* [online]. Message to: Ladislav CUBR. 2015-07-10 [cit. 2015-07-15]. Osobní komunikace.
- Finsko. Opetus- ja kulttuuriministeriö. 2014. National Digital Library of Finland standard portfolio. In *The National Digital Library* [online]. 27.10.2014, 14 s. [cit. 2015-06-10]. Dostupné z: http://www.kdk.fi/images/tiedostot/NDL_standardportfolio27_10_2014.pdf.
- Hakala, Juha. (2001a, květen). *Popis dokumentů a přístup k nim – nové výzvy*. CASLIN 2001, Beroun, Česká republika. Dostupné z: <http://old.stk.cz/hakala.pdf>.
- Hakala, Juha. 2001b. *Using national bibliography numbers as Uniform Resource Names* [online]. 2001 [cit. 2015-07-15]. Internetový standard RFC 3188. Dostupné z: <https://www.ietf.org/rfc/rfc3188.txt>.

ISO 14721:2012. *Space data and information transfer systems - Open archival information system (OAIS) - Reference model*. Geneva: International Organization for Standardization, 2012 (česky ČSN ISO 14721. *Systémy pro přenos dat a informací z kosmického prostoru - Otevřený archivační informační systém - Referenční model*. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, 2014. 111 s. Třídící znak 319620.

ISO 26324:2012. *Information and documentation: Digital object identifier system*. Geneva: International Organization for Standardization, 2012.

ISO 2108:2005. *Information and documentation: International standard book number (ISBN)*. Geneva: International Organization for Standardization, 2005.

ISO 3297:2007. *Information and documentation: International standard seriál number (ISSN)*. Geneva: International Organization for Standardization, 2007.

Kahn, Robert, Wilensky, Robert. 2006. A Framework for Distributed Digital Object Services. *International Journal on Digital Libraries*. 2006, Vol. 6, Issue 2, s. 115-123 [cit. 2015-07-15]. ISSN 1432-1300. Dostupné z: https://www.doi.org/topics/2006_05_02_Kahn_Framework.pdf.

Kansalliskirjasto. 2007. URN - Uniform Resource Name. In *The National Library of Finland* [online]. 2007 [cit. 2015-06-10] Dostupné z: <http://www.nationallibrary.fi/publishers/urn.html>.

Karvonen, Minna. 2015. *The National Digital Library* [online]. Message to: Iveta LODROVÁ. 2015-01-22 [cit. 2015-06-10]. Osobní komunikace.

Keskitalo, Esa-Pekka. 2015. *URN:NBN* [online]. Message to: Iveta LODROVÁ. 2015-06-10 [cit. 2015-06-20]. Osobní komunikace.

Nestor working group long-term preservation standards. 2009. *Catalogue of criteria for assessing the trustworthiness of PI systems: Draft for public comment* [online]. Göttingen: Niedersächsische Staats- und Universitätsbibliothek, 2009 [cit. 2015-09-01]. um:nbn:de:0008-20080710227. Dostupné z: http://files.d-nb.de/nesstor/materialien/nesstor_mat_13_en.pdf.

Pietarilla, Janne. 2015. *URN: components of system* [online]. Message to: Iveta LODROVÁ. 2015-07-22 [cit. 2015-07-23]. Osobní komunikace.

PILIN project (Persistent Identifier Linking Infrastructure): project closure report. Victoria (Austrálie): Monash University, 2007 [cit. 2015-05-14]. 102 s.

Pistorius, Vladimír. *Jak se dělá kniha: příručka pro nakladatele*. 3., dopl. a přeprac. vyd. Příbram: Pistorius & Olšanská, 2011. 280 s. ISBN 978-80-87053-50-8.

Piščanc, Jordan. (2013, červen). *National Bibliography Number Italia (NBN:IT) Project: A persistent identifier supporting national legal deposit for digital resources* [online]. CERN Workshop on Innovations in Scholarly Communication (OAI8), University of Geneva. Dostupné z: <https://indico.cern.ch/event/211600/session/7/contribution/42>.

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

Řehánek, Martin. 2015. *Softwarové nástroje pro podporu URN:NBN ve vybraných evropských zemích*. 23.2.2015 [cit. 2015-05-15]. Interní studie.

Schweizerische Nationalbibliothek. 2014. *e-Helvetica: URN-Handbuch: Zusammengetragen und erstellt durch die URN-Koordinationsstelle der Schweizerischen Nationalbibliothek*. Update April 2014 [cit. 2015-05-14]. Dostupné z: http://www.dnb.de/EN/Service/DigitaleDienste/URNService/urnservice_node.html.

Vašek, Zdeněk, Cubr, Ladislav a kol. 2014. *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:cz:nk-0027gj. Dostupné z: http://www.ndk.cz/archivace/Metodika_URNNBN_final_2.1a.pdf.

Voigt, Rita. 2007. Persistent identifiers and finnish institutional repositories. In *SlidePlayer* [online]. Helsinki University of Technology, 2007 [cit. 2015-06-10]. Dostupné z: <http://slideplayer.biz/slide/2651721/>.