



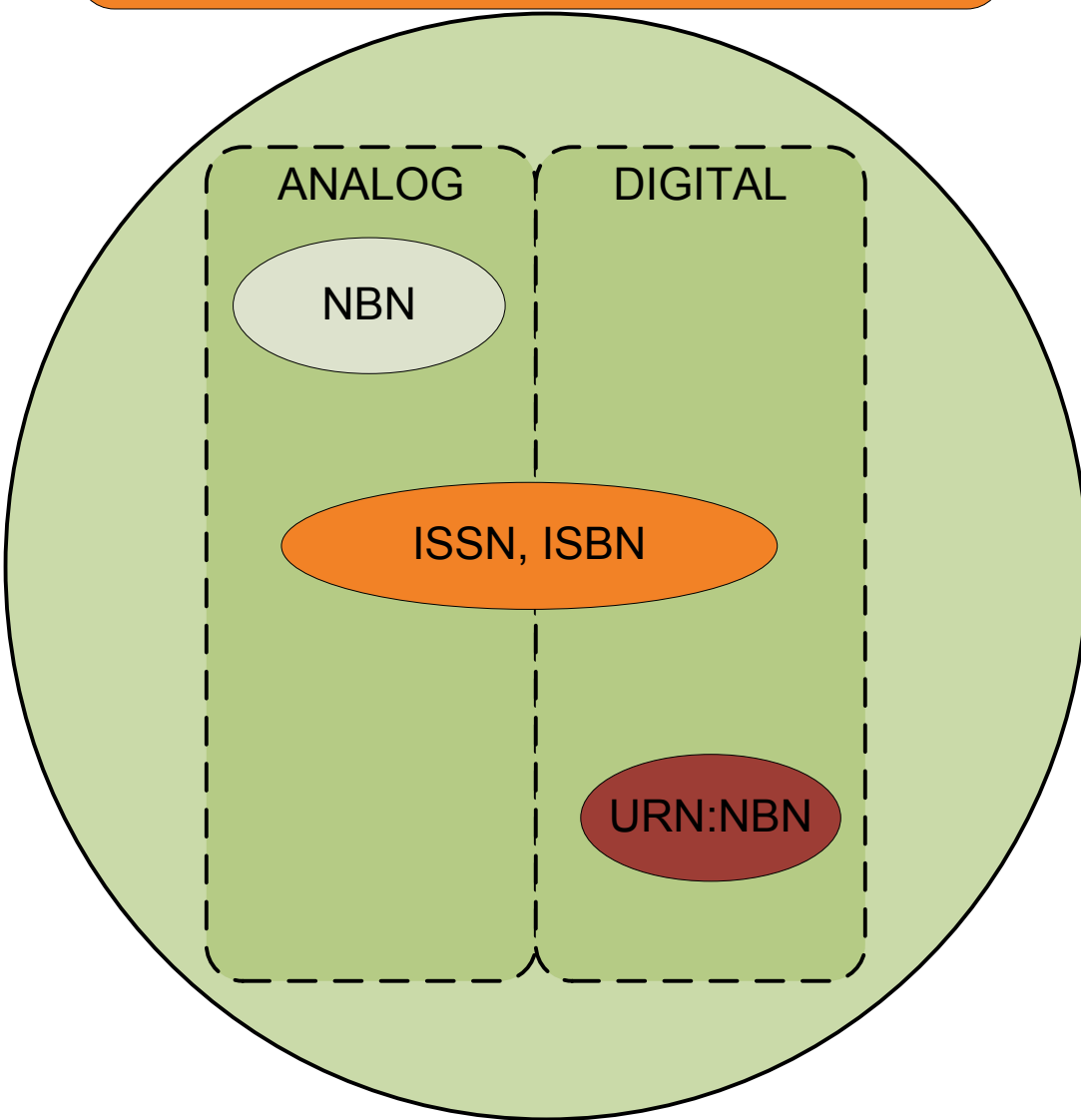
Identifikátory 2009

URN:NBN a číslo ČNB

Ladislav Cubr
Jan Hutař
Marek Melichar

NK ODO

Identifikátory a typy zdrojů, které mohou označovat - standard



National Bibliography Number (NBN) is a generic name referring to a group of identifier systems utilised by the national libraries and only by them for **identification of deposited publications which lack an identifier, or to descriptive metadata (cataloging) that describes the resources.**

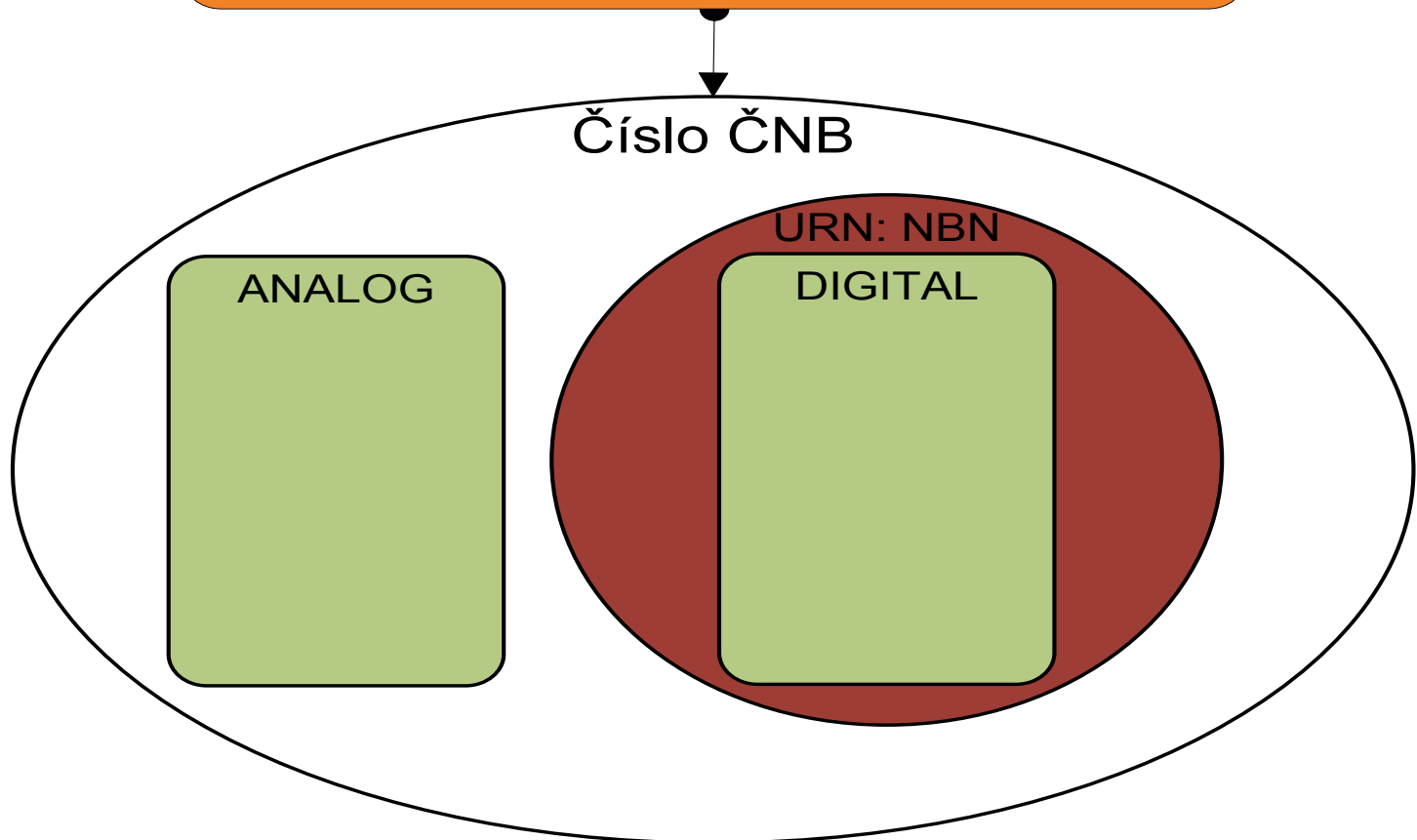
The materials identified by a national bibliography number may exist only in printed or other physical form, not electronically.

RFC3188

Pro náš kontext:

Identifikátory a typy zdrojů, které mohou označovat – aplikace pro CZ

Základní intelektuální entita (PREMIS) odpovídá 1 záznamu v katalogu ČNB





URN:NBN

Primární cíle URN:NBN

- Identifikace obsahu digitálních dokumentů a jejich aktuálních lokací
- Možnost získat přímo dokument – proto: dokumenty musí být přístupné online
- Dokumenty mají vztah k národní bibliografii a národnímu kulturnímu dědictví
- Pomůcka pro národní knihovny



Intelektuální entita

- množina informací, která je považována za jednu intelektuální jednotku pro účely správy a popisu, např. kniha [PREMIS, 2008]
 - 1 intelektuální entita může obsahovat jiné intelektuální entity (např. kniha „Tři mušketýři“ má svazky)
- je nutné stanovit *základní intelektuální entitu*, tj. základní jednotku pro identifikaci, popis a vyhledávání v našem kontextu

V kontextu NK by měla být základní IE ta entita, která odpovídá záznamu v bázi ČNB.



ČNB a FRBR

PRO TRADIČNÍ (ANALOGOVÉ) ZDROJE PLATÍ:

- záznam v ČNB je záznamem manifestace (*manifestation*) (FRBR, 2008)
- manifestace je „fyzické ztvárnění“ vyjádření díla
 - manifestace je množinou všech fyzických objektů, které nesou stejné charakteristiky co se týče intelektuálního obsahu a fyzické podoby
 - zahrnuje 1 nebo více fyzických objektů (např. 100 výtisků jednoho vydání knihy nebo 1 rukopis)
 - jednotka (*item*) je jedním exemplářem manifestace

PROTO: ČNB neobsahuje jednotky, ale manifestace (FRBR)



Digitální reprezentace

•Digitální reprezentace (*representation*) je množina všech počítačových souborů, které jsou potřebné k tomu, aby je mohl software reprodukovat jako jednu intelektuální entitu [PREMIS, 2008]

–např. 32 souborů ve formátu JPG2000 (32 digitalizovaných stránek tištěné knihy) a 1 soubor s metadaty > Máchův Máj (nakl. J.L.Kober, 1926)

!V kontextu digitalizovaných knih v NK se digitální reprezentace vždy váže k dané manifestaci, která je zaznamenána v bázi ČNB

!DR sama o sobě nemá vlastní číslo ČNB, ale „dědí“ jej z dané manifestace



Záznam ČNB a číslo ČNB (čČNB)

Záznam v ČNB obsahuje:

- bibliografický záznam IE (např. vydání Tří mušketýrů 1930)
 - případně také obsahuje informace o nižších IE, z nichž se základní intelektuální entita skládá (např. 3 svazky tohoto vydání)

Záznam by měl dále obsahovat:

- jedinečný a trvalý identifikátor ČNB (tj. číslo ČNB > čČNB)
 - číslo ČNB by mělo být od listopadu 2009
 - asi pole 015
 - **tím máme identifikaci manifestace**

Záznam by mohl dále obsahovat:

- seznam všech URN:NBN označujících digitální reprezentace dané manifestace, případně další identifikátory (např. ISBN)

Premis verze NKP	Příklad	Metadata	Číslo ČNB	URN:NBN	Handle
Základní intelektuální entita	Karel Čapek: Válka s mloky. Praha, 1936 (sysno 000743224)	záznam manifestace v katalogu ČNB	bude mít (např. 123456)	nebude mít	nebude mít
Digitální reprezentace 1	naskenovaná kniha (Městská knihovna)	bibliografické údaje z katalogu Koniáš (včetně ISBN)	má číslo ČNB 123456 v metadatech	URN:NBN:CZ:NDK:159753	může mít Handle v digitální knihovně MK
Počítačové soubory 1	1x pdf	metadata pro jednotlivé počítačové soubory	N/A	N/A	může mít Handle v digitální knihovně MK
Digitální reprezentace 2	naskenovaná kniha (Kramerius NK)	bibliografické údaje vybrané z ČNB (sysno 000743224)	má číslo ČNB 123456 v metadatech	URN:NBN:CZ:MLP:454545	má (digitální knihovna Kramerius)
Počítačové soubory 2	250x jpeg	metadata pro jednotlivé počítačové soubory	N/A	N/A	může mít (digitální knihovna Kramerius)

Konec první části

Po přestávce budeme pokračovat ;-)

identifikátor (URN:NBN) není
magický nástroj, který vyřeší
různé problémy různých
oblastí digitalizace



Obeční rámec

- 1) Potřeba nové konceptualizace a nového řešení – důvody:
 - větší komplexita než se očekávalo
 - „nedostatečnost“ řešení pomocí italského resolveru
 - nereálná očekávání
- 2) **potřeba** vytvořit konsorcium? zainteresovaných institucí (NK, STK, MZK aj.), kt. mohou nabídnout (lidi, znalosti)
- 3) konsorcium pak požádá o granty/financování (VISK3?)
- 4) konsorcium musí být „projektově řízené“ s jasným cílem a jasně vymezenými odpovědnostmi, zdroji financí...



Obeční rámec pokrač.

5) potřeba vyjasnit si znovu základní pojmy (pracovní návrh koncepce ODO NK bude k připomínkování)

- potřeba se shodnout a porozumět si

6) potřeba vyjasnit si vztahy mezi institucemi a možnosti spolupráce

- např. RD.CZ je projekt, tedy jak moc je jeho trvalé financování a pokračování zaručeno?



Italský resolver

- software a celý italský systém je stále ve fázi vývoje
- špatná komunikace (Italové nereagují často)
- nemáme zatím žádný manuál ani instalační dokumentace
- chybí podrobný popis implementace celého systému

- bude potřeba úprav pro české prostředí (tj. lze na tom stavět)



Tools

NBN research

Categories

- All (0)
- Inconsistent (0)
- Withdrawn (0)

System state

Refresh

NBN active: 0

NBN inactive: 0

Last harvest time form parent:

Last sync P2P node:

Nodes

Refresh Delete Save

<input type="checkbox"/> Node name	BaseURL	Organization	Referent
Type	Content		
Node name:	<input type="text"/>		
BaseURL:	<input type="text"/>		
Organization:	<input type="text"/>		
Referent:	<input type="text"/>		
E-mail:	<input type="text"/>		
Telephone number:	<input type="text"/>		

Tools

Nodes



Základní potřeba je proto **vytvořit vlastní české řešení (koncept + resolver)**

– získat grant nebo další prostředky navíc

To obnáší dvě věci:

- nová konceptualizace systému
- naprogramování jeho softwarového řešení (úprava?)



**Samotný trvalý identifikátor sám o sobě nic neřeší,
pokud k němu není připojena řada metadat:**

- informace o intelektuálním obsahu (bibliografické informace)
- informace o aktuálních umístěních (adresy URL)
- administrativní metadata (pod jakou instituci objekt spadá)
- informace o právech

je nutné trvale udržovat:

- **registr identifikátorů**
- **všech přidružených metadat**
- **trvale vše aktualizovat**



System URN:NBN v ČR by měl dvě větve:

- 1.jedna pro NK a její potřeby (NDK; ČNB)
- 2.větev mimo NK (dodržení standardů NK)



Větev pro NK

všechny digitální reprezentace IE v ČNB budou mít
URN:NBN

–autoritativní systém spojující trvalou identifikaci IE (záznam v ČNB)
s jejich digitálními reprezentacemi

zejména pro potřeby NDK (masová digitalizace)

2 odlišné systémy identifikace:

- systém čísel ČNB
- systém URN:NBN



Větev mimo NK

- stejný SW, různá práva, různé využití
- dig. objekty většinou bez papírové předlohy (nemají záznam v ČNB, např. NUŠL a univerzity)
- vymežit, do jaké míry se bude NK angažovat
- MZK, KNAV aj.



Resolver hledá ...

– hledání podle:

- URN:NBN
- čČNB
- možná i podle jiných metadat



Co resolver zobrazí?

- zadáme URN:NBN > resolver vrátí

- stránka, kde budou všechna metadata k URN:NBN,
 - název, autor, číslo ČNB
 - odkazy na lokace všech reprezentací (např. složené objekty nebo i jednoduché DO)

- zadáme čČNB > resolver vrátí

- všechna URN:NBN související s tím čČNB
- ideálně i URL pro každé URN:NBN
- všechna metadata



Proč to tak dlouho trvá?

- ukazuje se, že problém je mnohem složitější
 - ve všech zemích to řešili několik let
 - IOP
 - neshoda v terminologii
 - chybí IT podpora v NK
-
- Hakala: Nejste pozadu, všude se stále řeší komplexnější případy identifikace digitálních reprezentací



Spojit / nespojit URN:NBN s RD.CZ?

není to obvyklé ve světě

- RD.CZ si vystačí s čísly ČNB – ta identifikují tutéž manifestaci (tedy např. vydání Tří mušketýrů z 1930)
- jak dlouho bude existovat RD.cz?
- jak do toho vstoupí nový systém zakoupený v IOP?
- lépe přidělovat již při „výrobě“?
- lépe když to bude samostatná aplikace?



Proč nedat čČNB do URN:NBN syntaxe?

- technicky je to složitější a není to obvyklé
- obecně se to nedoporučuje
- pokud bude mít ČNB chyby (duplicity aj.), pak by bylo devalvováno i URN:NBN
- a hlavně – proč?
 - V metadatech identifikátoru ČNB bude, stejně jako další metadata...



Co chceme?

- vytvořit skupinu knihoven pro podání projektu
- dotáhnout konceptualizaci trvalé identifikace
- vytvořit podmínky pro zapojení dalších oddělení NK (katalogizátorů ap.)
- definovat koncepci a podle ní systém (resolver)

- nouzové řešení využití hotového (ovšem nedoladěného) SW odjinud (Italové apod.)