

Orientace ve standardech digitalizace

Studijní text

Kvalifikační úroveň 7

Verze 1.0

Autor: Mgr. Markéta Krutská

Konzultant: Mgr. Zdeněk Hruška

Ostrava 2021



Odborná způsobilost	Orientace ve standardech digitalizace
Profesní kvalifikace	Knihovník specialista správce digitální knihovny (kód: 72-025-T)
Kvalifikační úroveň	7
Klíčová slova	Identifikátory, FRBR, čárový kód, signatura, systémové číslo, ISBN, ISSN, ISMN, číslo České národní bibliografie, UUID, URI, URL, PURL, URN, URN:NBN, Handle, DOI, ARK, metadata, popisná metadata, administrativní metadata, technická metadata, archivační metadata, právní metadata, strukturální metadata, metadatové schéma, XML, XSD, MARC21, MARCXML, Dublin Core, MODS, PREMIS, MIX, TextMD, ALTO, METS, standardy NDK.
Anotace	Tento studijní text se zabývá identifikátory a metadaty, které se využívají při digitalizaci v knihovnách. Ve studijním textu je vysvětleno, k čemu identifikátory a metadata slouží a proč je potřeba je při digitalizaci používat. Následně jsou popsány konkrétní identifikátory a metadatové formáty, které se používají v mnoha českých knihovnách. Studijní text je určen pracovníkům knihoven a dalších paměťových institucí, kteří se chystají digitalizovat své fondy či se jen zajímají o tuto problematiku.

Tvorba studijních textů je realizována za finanční podpory Ministerstva kultury České republiky v rámci projektu Veřejné informační služby knihoven (VISK 1).

Zkouška z odborné způsobilosti **Orientace ve standardech digitalizace** má 4 části, zkoušený musí splnit tato kritéria hodnocení:

Kritéria hodnocení	Způsob ověření
a Charakterizovat typy identifikátorů dokumentů a digitálních objektů v knihovnách, jejich účel a využitelnost v digitálním prostředí	Ústní ověření
b Vytvořit vlastní metadatové schéma pro popis určitého typu dokumentu, a to včetně pravidel popisu a uvedení příkladu	Praktické předvedení a ústní ověření
c Orientovat se v dokumentech XML (Extensible Markup Language, Rozšiřitelný značkovací jazyk)	Praktické předvedení a ústní ověření
d Charakterizovat metadata, jejich funkce a typy a metadatové standardy	Praktické předvedení a ústní ověření

Tento text pro vás připravila Mgr. Markéta Krutská, Moravská zemská knihovna v Brně, marketa.krutska@mzk.cz

Text konzultoval Mgr. Zdeněk Hruška, Moravská zemská knihovna v Brně, zdenek.hruska@mzk.cz

OBSAH

RYCHLÝ NÁHLED STUDIJNÍHO TEXTU	6
1 IDENTIFIKÁTORY	7
1.1 Model FRBR – provedení vs. jednotka	7
1.2 Identifikátory používané v knihovnách.....	8
1.2.1 Čárový kód	8
1.2.2 Signatura	11
1.2.3 Systémové číslo (zkráceně sysno)	13
1.2.4 ISBN, ISSN, ISMN	14
1.2.5 Číslo České národní bibliografie (čČNB)	19
1.3 Identifikátory používané v digitálním prostředí	20
1.3.1 UUID	20
1.3.2 URI	21
1.3.3 URL	22
1.3.4 PURL	23
1.3.5 URN.....	24
1.3.6 URN:NBN	25
1.3.7 Handle	26
1.3.8 DOI	27
1.3.9 ARK	28
2 METADATA.....	29
2.1 Metadatové schéma	29
2.2 XML.....	31
2.2.1 XSD	34
2.3 Popisná metadata	35
2.3.1 MARC 21, MARCXML	35
2.3.2 Dublin Core.....	36
2.3.3 MODS.....	37
2.4 Administrativní metadata.....	39
2.4.1 PREMIS.....	39
2.4.2 MIX	40

2.4.3	TextMD.....	40
2.4.4	ALTO.....	40
2.5	Strukturální metadata	42
2.5.1	METS	42
2.6	Standardy Národní digitální knihovny (NDK)	43
2.6.1	Formátový výbor NDK.....	43
SHRNUTÍ STUDIJNÍHO TEXTU.....		45
LITERATURA		46

RYCHLÝ NÁHLED STUDIJNÍHO TEXTU

ÚVOD

Hlavní činností správce digitální knihovny je správa digitálních dokumentů uložených v digitální knihovně. Správce digitální knihovny se tak každodenně setkává s identifikátory a s metadaty digitálních dokumentů. Identifikátory se využívají v celém procesu digitalizace, tedy jak při skenování a úpravě skenů, tak při zpřístupňování a při archivaci dokumentů. Metadata se pak používají pro popis digitálních objektů. Je proto nutné, aby se správce digitální knihovny v této problematice orientoval, tedy aby chápal, proč jsou identifikátory a metadata důležité, a aby znal konkrétní systémy identifikátorů a metadatové standardy, které se při digitalizaci používají. Tento studijní text má za cíl všechny zmíněné znalosti pokrýt.

PO PROSTUDOVÁNÍ STUDIJNÍHO TEXTU BUDETE UMĚT

- vysvětlit, co to identifikátory jsou a k čemu slouží,
- popsat konkrétní identifikátory, které se používají při digitalizaci v knihovnách,
- vysvětlit, co to jsou metadata a k čemu slouží,
- popsat typy metadat a u každého typu znát konkrétní metadatové formáty,
- vytvořit vlastní metadatové schéma včetně pravidel popisu a uvedení příkladu,
- vysvětlit, co je to XML, a nalézt chyby v XML záznamu.

KLÍČOVÁ SLOVA STUDIJNÍ OPORY

Identifikátory, FRBR, čárový kód, signatura, systémové číslo, ISBN, ISSN, ISMN, číslo České národní bibliografie, UUID, URI, URL, PURL, URN, URN:NBN, Handle, DOI, ARK, metadata, popisná metadata, administrativní metadata, technická metadata, archivační metadata, právní metadata, strukturální metadata, metadatové schéma, XML, XSD, MARC21, MARCXML, Dublin Core, MODS, PREMIS, MIX, TextMD, ALTO, METS, standardy NDK.

1 IDENTIFIKÁTORY

Identifikátor je “specifický deskriptor označující jedinečnou entitu”.¹ **Pomocí identifikátorů se rozlišují jednotlivé entity v určitém systému.** V prostředí knihoven se používají různé typy identifikátorů, rozlišovat můžeme například dokumenty v knihovním fondu, čtenáře knihovny či záznamy v elektronickém katalogu. V digitálním prostředí jsou identifikátory též velmi potřebné, můžeme pomocí nich rozlišit samotnou digitální knihovnu, digitální dokumenty v digitální knihovně, strany v rámci jednoho dokumentu, články a kapitoly v dokumentu či virtuální sbírky dokumentů. Identifikátory se mohou skládat z číslic, písmen a dalších znaků. V následujících kapitolách se budeme zabývat konkrétními identifikátory, které se používají v knihovnách a které jsou důležité pro digitalizaci a archivaci dokumentů.

1.1 Model FRBR – provedení vs. jednotka

Než se začneme zabývat jednotlivými identifikátory, měli bychom si nejprve sjednotit terminologii, která se používá pro pojmenování objektů popisovaných v knihovnách. V roce 1997 publikovala IFLA (International Federation of Library Associations and Institutions = Mezinárodní federace knihovnických sdružení a institucí) studii „Funkční požadavky na bibliografické záznamy“² (Functional Requirements for Bibliographic Records = FRBR). Cílem této studie bylo **mezinárodní sjednocení ve vytváření a v chápání bibliografických záznamů**. Stěžejní část studie tvoří popis modelu, kterému se podle názvu studie říká FRBR model. Model popisuje entity (klíčové objekty zájmu uživatelů), atributy (vlastnosti entit) a vztahy mezi entitami. Pro nás jsou důležité entity dílo, vyjádření, provedení a jednotka.

- **dílo** = Určitý intelektuální nebo umělecký výtvor. Jedná se o abstraktní entitu, neexistuje žádný materiální objekt, který by se dal označit jako dílo.
- **vyjádření** = Určitá intelektuální nebo umělecká realizace díla ve formě abecedně-číselného, hudebního nebo choreografického zápisu, zvukového záznamu, obrazového statického nebo pohyblivého záznamu, objektu atd., nebo kombinace těchto forem.
 - Tzn. jedno dílo může mít podobu textu, divadelní hry a filmu, a pak se jedná o tři různá vyjádření jednoho díla.
- **provedení** = Fyzická podoba vyjádření díla. Všechny hmotné objekty, které mají stejné vlastnosti, pokud jde o intelektuální obsah a fyzickou formu.

¹ BALÍKOVÁ, Marie. Identifikátor (indexace). In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2021-10-01]. Dostupné z: https://aleph.nkp.cz/F/?func=direct&doc_number=000001548&local_base=KTD

² Pokud by Vás studie FRBR zajímala více, [zde](#) je přeložená do češtiny.

Jakákoliv změna formy (změna formátu, písma, přebalu, papíru) znamená nové provedení.

- Např. druhé vydání knihy v nakladatelství XYZ je jedno provedení, třetí vydání knihy v témže nakladatelství je už další provedení.
- **jednotka** = Jednotlivý exemplář provedení, tedy konkrétní entita.
 - Např. když se jedno vydání knihy (provedení) vytiskne ve více kopiích, tak jednotlivé kopie nazýváme jednotky.

V praxi se nejčastěji setkáme s identifikátory, které popisují provedení (vydání) a jednotky (konkrétní exempláře). Například pokud se rozhodneme, že zdigitalizujeme Kytici od K. J. Erbena, musíme si nejprve vybrat, jaké provedení zdigitalizujeme. Kytice vyšla v mnoha vydáních a my musíme umět ta vydání od sebe rozlišit. Může se stát, že knihovna vlastní více kopií jednoho vydání, tedy vlastní více jednotek stejného provedení. I jednotky je potřeba od sebe odlišit, takže i jednotky mívají svůj identifikátor.

1.2 Identifikátory používané v knihovnách

Nyní si představíme identifikátory, které se používají pro rozlišení dokumentů v knihovnách. Tyto identifikátory jsou důležité i pro digitalizaci, jelikož slouží k provázání fyzických dokumentů s digitálními kopiemi.

1.2.1 ČÁROVÝ KÓD

Čárové kódy slouží k **jednoznačnému rozlišení knihovních jednotek (exemplářů) v knihovním fondu**. Systém využití čárových kódů v českých (a slovenských) knihovnách vznikl v 80. letech 20. století kvůli automatizované evidenci výpůjček. Pokud knihovna vlastní více knihovních jednotek stejného provedení díla, tak ke každé jednotce je přidělen jiný čárový kód. V rámci jednoho knihovního fondu by měly být čárové kódy jedinečné. Čárové kódy lze snímat čtečkou, takže odpadá zdlouhavé ruční přepisování, při kterém by se mohly udělat chyby.

Čárové kódy nejsou tvořeny zcela náhodnými číslicemi. Na začátku čárového kódu se nachází prefix, který označuje knihovnu, které knihovní jednotka patří. Prefixy přiděluje Národní technická knihovna.³

³ Na webu Národní technické knihovny se nachází podrobné informace k pravidlům vytváření čárových kódů a seznam knihoven s přidělenými prefixy. <https://www.techlib.cz/cs/2920-carove-kody-pro-ceske-knihovny>

Čárový kód má vždy 10 číslic a skládá se ze tří prvků:

1. první prvek (jeden znak) vyjadřuje počet znaků použitých pro označení knihovny a zároveň signalizuje velikost knihovny,
2. následuje číselný kód knihovny (1 až 4 znaky), vyjadřuje tedy lokaci knihovní jednotky,
3. a konečně vlastní označení konkrétní knihovní jednotky (8 až 5 znaků) v rámci dané knihovny.

Velikost fondu knihovny (řádově)	Struktura kódu
desítky milionů	1xyyyyyyyy
miliony	2xxyyyyyyy
statisíce	3xxxxyyyyy
desetitisíce	4xxxxxyyyy

Obrázek 1 – x jsou číslice vyjadřující označení knihovny, y jsou číslice vyjadřující označení knihovní jednotky

Čárové kódy se používají v procesu digitalizace pro jednoznačné rozlišení digitalizovaných knihovních jednotek a pro provázání knihovního záznamu s digitálním dokumentem. I po digitalizaci je užitečné mít čárový kód uvedený v metadatech digitálního dokumentu, může se to hodit pro různé opravy — například pokud se změní údaje v knihovním záznamu, tak pomocí čárového kódu můžeme nalézt digitální dokument a aktualizovat popisné údaje (nebo naopak můžeme objevit chybu v digitálním dokumentu a pomocí čárového kódu dohledáme knihovní záznam použitý k popisu a opravíme ho). Čárové kódy také mohou sloužit k vyhledávání digitálních dokumentů v dlouhodobém úložišti.



Obrázek 2 – Zde jsou ukázky čárových kódů. I kdyby nad kódy nebylo napsané, jaké instituci dokument patří, dalo by se to odvodit a dohledat na webu Národní technické knihovny. Např. u čárového kódu 3142262036 nám první znak říká, že pro označení instituce jsou použity tři znaky, v tomto případě 142 (číslovka 3 nám též říká, že velikost fondu je v řádu stotisíců). Na webu NTK si můžeme dohledat, že prefix 3142 patří právě knihovně VŠE. Zbytek čárového kódu 262036 je označení dané knihovní jednotky.

MZK Brno(BIB) - Knihovní jednotky

Čapek, Karel, 1890-1938
Zahradníkův rok / -- 1. vyd.
Praha : Studio trnka, 2009 -- 108 s. : ISBN: 978-80-87209-32-5

Budete-li chtít zadat požadavek na výpůjčku nebo na kopii jednotky, klepněte na odkaz "požadavek" nebo "kopie" (musíte se přihlásit, případně odkaz "kopie" je určen pro externí knihovny žádající o kopii v rámci MVS).

Zvolte rok Zvolte ročník Zvolte dílčí knihovnu ☐ Skrýt vypůjčené jednotky

[< Předchozí stránka](#)

	Status	Půjčeno do:	Dílčí knih.	Sbírka / Doba vyhledání	Signatura	Signatura-2	Popis	Poznámka	Čárový kód
Kopie Detaily	Jen do studovny	prezenčně	MZK	Sklad / do 1 hodiny	2-1245.618				2610443351
Kopie Detaily	Měsíc	volné	MZK	Sklad / do 1 hodiny	2-1245.618				2610443352

Obrázek 3 – U jednoho provedení (konkrétní vydání díla) může knihovna vlastnit více knihovních jednotek a každá jednotka má svůj jedinečný čárový kód

1.2.2 SIGNATURA

Signatury **slouží k lokalizaci dokumentu v knihovním fondu**. Systémy signatur se mohou v různých knihovnách lišit (existují různé klasifikační systémy⁴ a knihovny si též mohou vytvořit vlastní nebo si upravit již existující systém). Více knihovních jednotek může mít stejnou signaturu a jedna jednotka může mít více signatur.

V rámci digitalizace je signatura důležitá hlavně pro vyhledání fyzického dokumentu, který se má digitalizovat. Později se se signaturou již moc nepracuje. Signatury se pak mohou zobrazovat u dokumentů v digitální knihovně a tato informace může být užitečná jak pro čtenáře, tak pro knihovníky.



Obrázek 4 – Každá knihovna může mít svůj vlastní systém signatur. Zde jsou ukázky, jak mohou signatury vypadat.



Obrázek 5 – Signatury se většinou uvádí na obálce knihy, aby byla na první pohled patrná lokace knihy

⁴ Mezi nejznámější a nejpoužívanější klasifikační systémy patří [Mezinárodní desetinné třídění](#) či [Třídění Kongresové knihovny](#).



Obrázek 6 – V digitální knihovně MZK se signatura zobrazuje u popisu knihy

MZK Brno(BIB) - Knihovní jednotky

Čapek, Karel, 1890-1938
 Zahradníkův rok / -- 1. vyd.
 Praha : Studio trnka, 2009 -- 108 s. : ISBN: 978-80-87209-32-5

Budete-li chtít zadat požadavek na výpůjčku nebo na kopii jednotky, klepněte na odkaz "požadavek" nebo "kopie" (musíte se přihlásit, případně Odkaz "kopie" je určen pro externí knihovny žádající o kopii v rámci MVS).

Zvolte rok Zvolte ročník Zvolte dílčí knihovnu ☐ Skrýt vypůjčené jednotky

< Předchozí stránka

	Status	Půjčeno do:	Dílčí knih.	Sbírka / Doba vyhledání	Signatura	Signatura-2	Popis	Poznámka
Kopie Detaily	Jen do studovny	prezenčně	MZK	Sklad / do 24 hodin	2-1245.618			
Kopie Detaily	Měsíc	volně	MZK	Sklad / do 24 hodin	2-1245.618			

Obrázek 7 – Více knihovních jednotek může mít stejnou signaturu. Na regálu v knihovně jsou dokumenty vedle sebe. Knihovní jednotky jsou od sebe rozlišeny čárovým kódem.

Hendl, Jan, 1947-
 Kvalitativní výzkum : -- Čtvrté, přepracované a rozšířené vydání
 Praha : Portál, 2016 -- 437 stran : ISBN: 978-80-262-0982-9

Budete-li chtít zadat požadavek na výpůjčku nebo na kopii jednotky, klepněte na odkaz "požadavek" nebo "kopie" (musíte se přihlásit, případně Odkaz "kopie" je určen pro externí knihovny žádající o kopii v rámci MVS).

Zvolte rok Zvolte ročník Zvolte dílčí knihovnu ☐ Skrýt vypůjčené jednotky

< Předchozí stránka

	Status	Půjčeno do:	Dílčí knih.	Sbírka / Doba vyhledání	Signatura	Signatura-2	Popis	Poznámka	Čárový kód
Kopie Detaily	Vol.výb.-měs.	16/07/2021	Vol.výběr	Absenční volný výběr - přízemí	2-1362.217	3 HEN			2610663850
Kopie Detaily	Vol.výb.-měs.	volně	Vol.výběr	Absenční volný výběr - přízemí	2-1362.217	3 HEN			2610663851
Kopie Detaily	Vol.výb.-měs.	28/06/2021	Vol.výběr	Absenční volný výběr - přízemí	2-1362.217	3 HEN			2610663852
Kopie Detaily	Vol.výb.-měs.	07/07/2021	Vol.výběr	Absenční volný výběr - přízemí	2-1362.217	3 HEN			2610688166

Obrázek 8 – Jedna knihovní jednotka může mít více signatur. Například v Moravské zemské knihovně se používají dva systémy signatur – signatury pro sklady, které jsou založené na formátu dokumentu, a signatury pro studovny, které jsou založené na věcném popisu a příjmení autora (tento systém je přívětivější pro čtenáře).

1.2.3 SYSTÉMOVÉ ČÍSLO (ZKRÁCENĚ SYSNO)

Každý záznam v online knihovním katalogu má svůj identifikátor, kterému se zkráceně říká sysno. Na jednom knihovním záznamu může být více knihovních jednotek, takže více jednotek může mít stejné sysno (a jak už víme, knihovní jednotky mohou mít různé signatury, takže pod jedním záznamem/sysnem může být více signatur.) Systémové číslo je jedinečné pouze v rámci dané knihovny, jiné knihovny mají svůj vlastní systém sysen.

Častou praxí při digitalizaci je to, že ke skenům se přiřazují záznamy z katalogu. Pak totiž není potřeba přepisovat popisné údaje (název, autor, nakladatel, ISBN atd.), což ušetří mnoho času a předejde se chybám. Sysno může být zaznamenáno v metadatech digitálních dokumentů, což je velmi užitečné, protože pak je snadno dohledatelné, jaký konkrétní záznam byl ke skenům přiřazen (může se to hodit pro různé opravy popisných údajů).

Signatura	→ 2-1245.618
Autoři	→ Čapek, Karel, 1890-1938
Název	→ Zahradníkův rok /
Autoři,ost.	→ Redouté, Pierre-Joseph, 1759-1840
Vydání	1. vyd.
Místo vyd.	→ Praha :
Nakladatel	→ Studio tmka.
Rok vydání	→ 2009
Rozsah	108 s. : il. (převážně barev.) ; 23 cm
Edice	→ Edice naší babičky
ISBN	978-80-87209-32-5 978-80-87209-33-2
ČNB	cnb001997242
Poznámka	Ilustrace: Joseph-Pierre Redouté Část nákladu s přílohou: Bylinkový kalendář báby kořenářky : byliny, herbář, zdraví, recepty ISBN sady s kalendářem jen na PVC přebalu Dostupné též v digitální podobě v aplikaci Kramerius
Předmět.heslo	→ zahrádkářství
Forma a žánr	* české prózy * fejetony * kalendáře
Exempláře	Všechny jednotky
Externí odkaz	http://www.digitalniknihovna.cz/mzk/uuid/uuid:b4bcc350-44f7-11e7-a34b-005056827e51
Sysno	001104126

Obrázek 9 – Ukázka sysna v katalogu MZK

NKC - Úplné zobrazení záznamu

Zvolte formát: Standardní -- Katalogizační záznam -- Stručný záznam -- MARC -- Citace

Záznam 1 z 1

Hlavní záhlaví	● Kafka, Franz, 1883-1924
Název	● Schloss, Francouzsky
Název	● Le chateau = (Das Schloss) / Franz Kafka ; Traduit de l'allemand par Alex
Nakl. údaje	Paris : Gallimard, 1938
Popis (rozsah)	255 s. ; 8°
EXEMPLÁŘE	
Forma, žánr	* německy psané romány
Skupina Konspektu	● 821.112.2-3 - Německá próza, německy psaná
MDT	● 821.112.2(437.3)-31 ● (0:82-31)
Systém. číslo	001104126

[https://aleph.nkp.cz/F/?func=direct&doc_number=001104126&local_base=NKC]

Obrázek 10 – Porovnejte tento snímek s předchozím. Zde máme záznam z katalogu Národní knihovny ČR. Sysno je stejné, ale jedná se o úplně jiný dokument. Sysno je unikátní pouze v rámci jedné instituce.

1.2.4 ISBN, ISSN, ISMN

Jedná se o **mezinárodní identifikátory, které slouží k jednoznačné identifikaci jednotlivých vydání (provedení) dokumentů**. Vydavatelé nejsou povinni se do těchto systémů zapojit, avšak používání těchto identifikátorů zjednodušuje identifikaci dokumentů jak vydavatelům, tak i prodejcům a knihovnám.

Přímo při digitalizaci se tyto identifikátory nepoužívají, nicméně opět je užitečné mít je zaznamenané v metadatech. Tyto identifikátory je vhodné zobrazovat i v digitální knihovně, aby čtenáři a knihovníci mohli jednoznačně identifikovat požadovaný dokument.

ISBN

ISBN (International Standard Book Number = mezinárodní standardní číslo knihy) je identifikátor sloužící k **rozlišení jednotlivých vydání knih**. V České republice byl tento systém zaveden v roce 1989, starším publikacím se ISBN zpětně nepřiděluje. Vrcholným celosvětovým orgánem systému je [Mezinárodní agentura ISBN](#) v Londýně, naším vrcholným orgánem je [Národní agentura ISBN v ČR](#) pracující při Národní knihovně České republiky.

Původně se ISBN skládalo z 10 čísel, od roku 2007 se přidělují ISBN s 13 čísly (přidal se prefix). Přechod na 13 znaků zajistil rozšíření kapacity systému a zároveň se nyní ISBN rovná EAN kódu (EAN kód s 13 znaky se celosvětově používá pro značení zboží).

ISBN má pevně danou strukturu a z jednotlivých částí identifikátoru lze získat základní informace o původu knihy. ISBN se skládá z pěti částí oddělených spojovníkem, první a poslední část mají pevnou délku, další tři části mají délku proměnlivou.

1. První částí je prefix. V současné době se používají dva prefixy – 978 a 979.
2. Druhou částí je tzv. identifikátor skupiny, který popisuje zemi, případně zeměpisný či jazykový region. Tato část se může skládat z jedné až pěti cifer. Identifikátory skupiny přiděluje Mezinárodní agentura ISBN. Například anglicky mluvící země mají kód 0 a 1, francouzsky mluvící země kód 2, německy mluvící země kód 3, atd. Pro Česko a Slovensko se používá kód 80.
3. Ve třetí části se nachází identifikace vydavatele. Tato část se může skládat z jedné až sedmi cifer. Každý vydavatel má svůj kód, kódy přidělují jednotlivé regionální agentury (v ČR to je již zmíněná Národní agentura ISBN fungující při Národní knihovně ČR).
4. Čtvrtou částí je kód pro konkrétní vydání knihy u daného vydavatele. Toto číslo může mít až šest cifer; kvůli pevné délce může být zarovnáno zleva nulami. Kód si určují sami vydavatelé.
5. Poslední částí je kontrolní číslice, která má vždy jeden znak.⁵

⁵ Popis výpočtu kontrolní číslice je popsán [zde](#)

Např. 978-80-200-3129-7

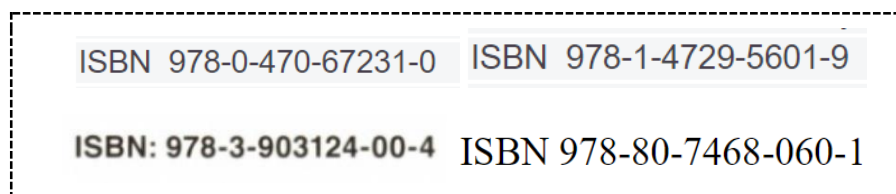
978 = prefix

80 = kód pro ČR

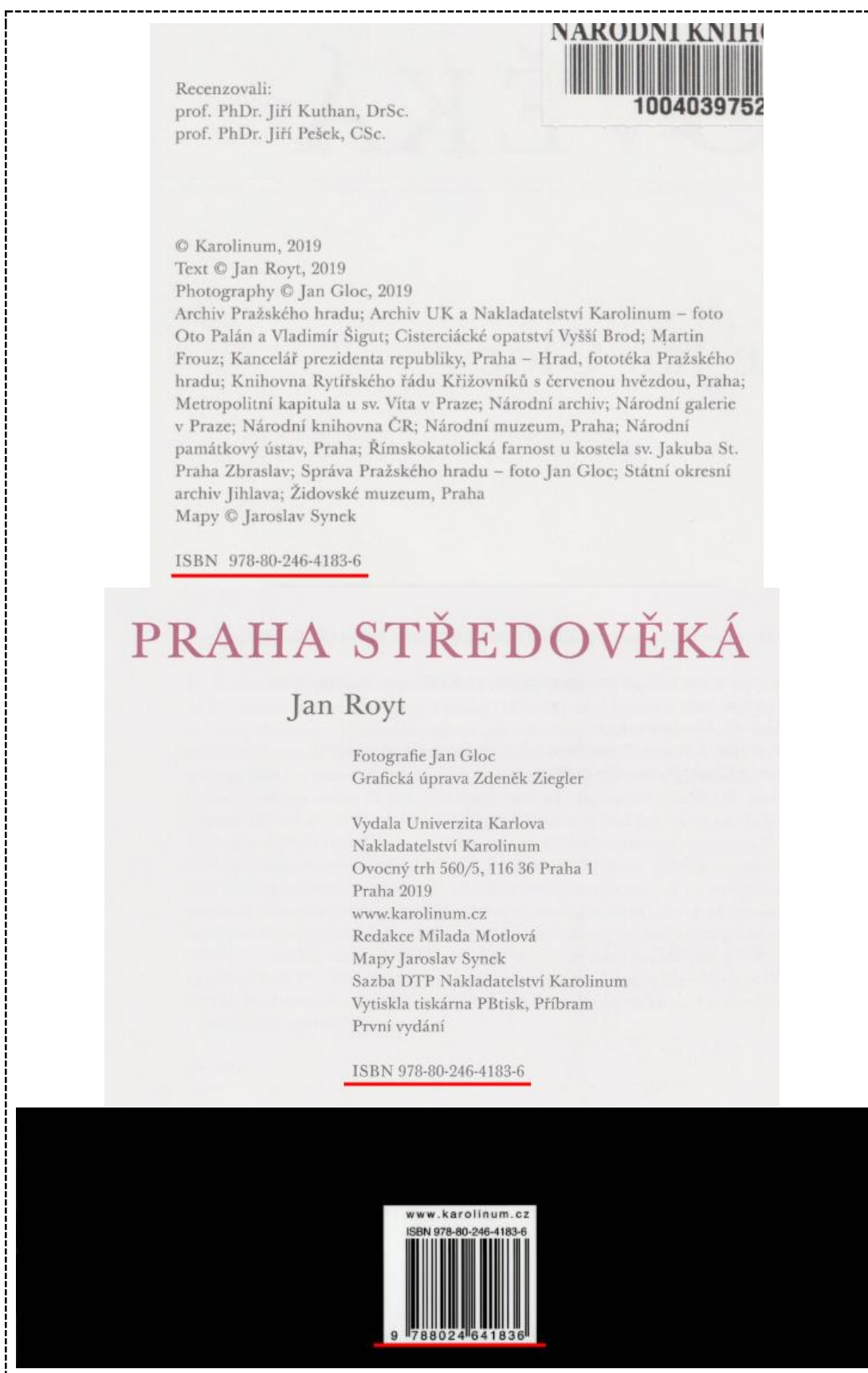
200 = identifikace vydavatele, v tomto případě se jedná o nakladatelství Academia

3129 = kód pro konkrétní vydání knihy

7 = kontrolní číslice



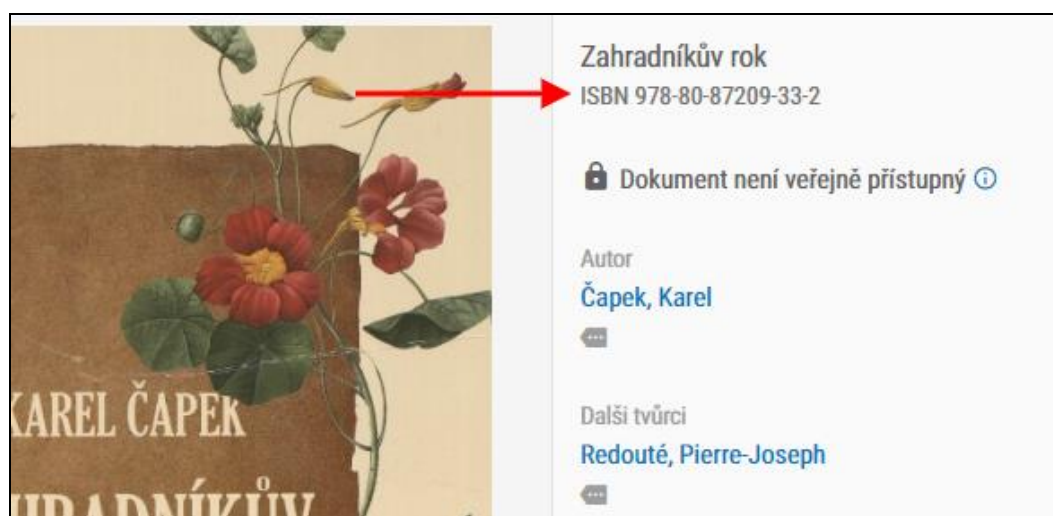
Obrázek 11 – Ukázky ISBN



Obrázek 12 – V knihách vydaných v ČR bývá ISBN zpravidla uvedeno na rubu titulního listu, v tiráži a na zadní obálce ve formě EAN kódu. U výše vyobrazené knihy se ISBN nachází na všech třech těchto místech.

Signatura	→ 2-1245.618
Autoři	→ Čapek, Karel, 1890-1938
Název	→ Zahradníkův rok /
Autoři, ost.	→ Redouté, Pierre-Joseph, 1759-1840
Vydání	1. vyd.
Místo vyd.	→ Praha :
Nakladatel	→ Studio trnka,
Rok vydání	→ 2009
Rozsah	108 s. : il. (převážně barev.) ; 23 cm
Edice	→ Edice naší babičky
ISBN	978-80-87209-32-5 978-80-87209-33-2
ČNB	cnb001997242
Poznámka	Ilustrace: Joseph-Pierre Redouté Část nákladu s přílohou: Bylinkový kalendář báby kořenářky : byliny, herbář, zdraví, recepty, ISBN sady s kalendářem jen na PVC přebalu Dostupné též v digitální podobě v aplikaci Kramerius
Předmět.heslo	→ zahrádkářství
Forma a žánr	* české prózy * fejetony * kalendáře
Exempláře	Všechny jednotky
Externí odkaz	http://www.digitalniknihovna.cz/mzk/uuid/uuid:b4bcc350-44f7-11e7-a34b-005056827e51 - D
Sysno	001104126

Obrázek 13 – Ukázka ISBN v katalogu MZK. V tomto případě jsou zde uvedené dva identifikátory ISBN, protože kniha vycházela samostatně a pak ještě s kalendářem a každá verze má své ISBN. ISBN se pak často ještě liší u brožované (paperback) verze, u vázané verze a u e-knih (u e-knih se ISBN liší i u různých formátů – např. EPUB, PDF, atd.). Většinou se totiž liší prodejní ceny, a pokud by bylo přiděleno pouze jedno ISBN, prodejce by měl problém s rozlišením.



Obrázek 14 – V digitální knihovně MZK se ISBN zobrazuje u popisu knihy, pomáhá to knihovníkům a čtenářům jednoznačně identifikovat dokument

ISSN

ISSN (International Standard Serial Number = mezinárodní standardní číslo seriálové publikace) slouží k **jednoznačné identifikaci pokračujících zdrojů (noviny, časopisy, sborníky, výroční publikace atd.)**. ISSN je přidělováno tištěným i online publikacím. Číslo ISSN se vztahuje k tzv. klíčovému názvu publikace, takže se změnou názvu se mění ISSN. Některé knihy, které jsou součástí série, mají přidělené ISBN i ISSN (jedno ISSN pro celou sérii a pak jedinečné ISBN pro každý svazek). V České republice byl tento systém zaveden na začátku 70. let, starším publikacím

se ISSN přiděluje i zpětně (na rozdíl od ISBN). Vrcholným celosvětovým orgánem systému je [Mezinárodní centrum ISSN v Paříži](#), naším vrcholným orgánem je [České národní středisko ISSN](#), které sídlí v Národní technické knihovně v Praze.

ISSN se skládá z 8 číslic, respektive ze dvou skupin číslic po čtyřech číslech. Dvě skupiny číslic jsou spojené spojovníkem. Poslední číslice je kontrolní a může nabývat hodnot 0–9 nebo X (s významem 10).⁶ Na rozdíl od ISBN se z ISSN nedají zjistit základní informace o původu publikace.

Např. 1802-0887

Národní technická knihovna spravuje národní databázi ISSN, která je veřejně přístupná: http://aleph.ntkcz.cz/F/?func=find-b-0&local_base=stk02



Obrázek 15 – Stejně jako ISBN i ISSN se může zobrazovat v digitální knihovně

ISMN

ISMN (International Standard Music Number = mezinárodní standardní číslo hudebnin) slouží k **jednoznačné identifikaci tištěných hudebnin**. V České republice bylo ISMN zavedeno v roce 1996. Vrcholným celosvětovým orgánem systému je [Mezinárodní agentura ISMN v Berlíně](#). Naším vrcholným orgánem je [Národní agentura ISMN v ČR](#) pracující při Národní knihovně ČR.

Původně se ISMN skládalo z 10 znaků a prvním znakem bylo písmeno M (např. M-2306-7118-7). Od roku 2008 se přešlo na 13 znaků a ISMN nyní začíná prefixem 979-0. K existujícím ISMN se přidal prefix 979 a písmeno M se změnilo na nulu (např. 979-0-2306-7118-7). Důvodem přechodu na 13 znaků bylo stejně jako u ISBN rozšíření kapacity a přechod na EAN kódy. Po prefixu 979-0 následuje 9 číslic rozdělených do 3 částí, všechny části jsou rozdělené spojovníky.

⁶ Více informací ke kontrolnímu číslu se nachází [zde](#)

1.2.5 ČÍSLO ČESKÉ NÁRODNÍ BIBLIOGRAFIE (ČČNB)

Tento identifikátor se zavedl v roce 2010 pro potřeby digitalizace. Hlavním účelem tohoto identifikátoru je **zamezení duplicitní digitalizace**. Pokud se nějaká instituce chystá digitalizovat své dokumenty, měla by si nejprve zkontrolovat, zda dokumenty nebyly již digitalizovány či zda jiné instituce neplánují jejich digitalizaci. Tato kontrola se provádí pomocí [Registru digitalizace](#). Nejspolehlivější způsob kontroly je porovnání jedinečných identifikátorů dokumentů (na úrovni provedení, tedy jednotlivých vydání). Těmito identifikátory jsou například výše zmíněné ISBN, ISSN a ISMN. Problém je v tom, že tyto identifikátory jsou přidělené pouze novějším dokumentům a starší dokumenty často nemají žádný identifikátor, který by byl jedinečný napříč knihovnami. Proto se zavedlo čČNB, které **se přiděluje publikovaným dokumentům vydaným na území ČR od r. 1801 do současnosti**.⁷ čČNB je přidělováno Národní knihovnou ČR, která spravuje [bázi ČNB](#).

Jak již bylo zmíněno, v rámci přípravy digitalizace je třeba ověřit, zda dokument nedigitalizovala jiná knihovna. Pokud dokument zdigitalizovaný není a my se ho tedy chystáme zpracovat, měli bychom se ujistit, zda je k dokumentu přiřazeno čČNB. V případě, že čČNB přidělené není, bychom měli o přidělení požádat. Více informací k žádostem o přidělení čČNB se nachází [zde](#). Informaci o plánované či provedené digitalizaci bychom pak měli zanést i do Registru digitalizace.

CNB - Úplné zobrazení záznamu	
Zvolte formát: Standardní -- Katalogizační záznam -- Stručný záznam -- MARC -- Citace	
Záznam 1 z 1	
Hlavní záhlaví	Čapek, Karel, 1890-1938
Název	Zahradníkův rok / Karel Čapek
Vydání	1. vyd.
Nakl. údaje	Praha : Studio trnka, 2009
Popis (rozsah)	108 s. : il. (převážně barev.) ; 23 cm
	Digitalizovaný dokument
EXEMPLÁŘE	
Číslo nár.bibl.	cnb001997242
ISBN	978-80-87209-32-5 (váz.) 978-80-87209-33-2 (váz. + kalendář)
Edice	Edice naší babičky
Další původce	Redouté, Pierre-Joseph, 1759-1840
Poznámka	Ilustrace: Joseph-Pierre Redouté Část nákladu s přílohou: Bylinkový kalendář báby kořenářky : byliny, herbář, zdraví, recepty, zahrada 2010 ISBN sady s kalendářem jen na PVC přebalu
Předmět. heslo	zahrádkářství
Forma, žánr	* české prózy * fejetony * kalendáře
Skupina Konspektu	821.162.3-3 - Česká próza
MDT	821.162.3-3 634/635 (0.82-3) (0.82-4) (059)
Předm. hesla-angl.	* gardening

Obrázek 16 – Takto vypadá záznam v bázi ČNB. Červeně podtržené je číslo ČNB.

⁷ čČNB se nepřiděluje šedé literatuře, což jsou nepublikované a polopublikované dokumenty, které nejsou běžně dostupné prostřednictvím knižního trhu (např. výroční zprávy institucí, katalogy a programy veletrhů a výstav, divadelní programy, propagační materiály, pozvánky, nepublikované vysokoškolské práce). ČČNB se také nepřiděluje jednotlivým svazkům seriálu a jednotlivým svazkům vícesvazkového díla bez významných názvů částí.

1.3 Identifikátory používané v digitálním prostředí

I v digitálním prostředí je potřeba umět rozlišit jednotlivé entity. Identifikátory v digitálním prostředí však často mívají ještě jednu funkci, a to zprostředkování okamžitého přístupu k dokumentu. Knihovny a další paměťové instituce by měly usilovat o to, aby zajistily trvalou dostupnost digitálních dokumentů. K tomu slouží trvalé identifikátory. Při zavádění trvalých identifikátorů je třeba zvážit následující kritéria:⁸

- **globální jedinečnost** – Identifikátor by měl být unikátní nejen na úrovni instituce, ale pokud možno celosvětově. Nejvhodnější je využít nějaký celosvětově používaný systém, který je spravovaný národní nebo mezinárodní autoritou.
- **trvalost** – Trvalost se vztahuje k době existence identifikátoru. Identifikátor by nemělo jít zrušit a měl by fungovat i poté, co již neexistuje autorita, která ho přidělila.
- **zpřístupnitelnost** – Identifikátor by měl umožňovat i zpřístupnění zdroje, ne jen pouhou identifikaci.
- **spolehlivost** – Infrastruktura, na které funguje systém trvalé identifikace, by měla být neustále v provozu.
- **spravující autorita** – Mělo by se jednat o důvěryhodnou instituci, která bude schopná zajistit trvalost systému.
- **pružnost** – Systém by měl být rozšiřitelný tak, aby bylo možné do budoucna zahrnout nové typy dokumentů.
- **interoperabilita** – Do systému by mělo být snadné přistupovat. Lze toho docílit používáním otevřených standardů.
- **nákladnost** – Je potřeba zvážit nákladnost používání systému trvalé identifikace.

1.3.1 UUID

UUID (Universally Unique Identifier = Univerzální unikátní identifikátor) **slouží k identifikaci digitálních objektů**. Identifikátory UUID nejsou spravovány centrální institucí, takže UUID si může vytvořit kdokoli. Pokud je UUID generováno podle definovaných pravidel, je velmi vysoká pravděpodobnost, že UUID bude jedinečné. Čistě teoreticky se sice může stát, že se vygenerují dva stejné identifikátory, ale vzhledem k délce identifikátoru a k algoritmu generování je ta pravděpodobnost velmi nízká.

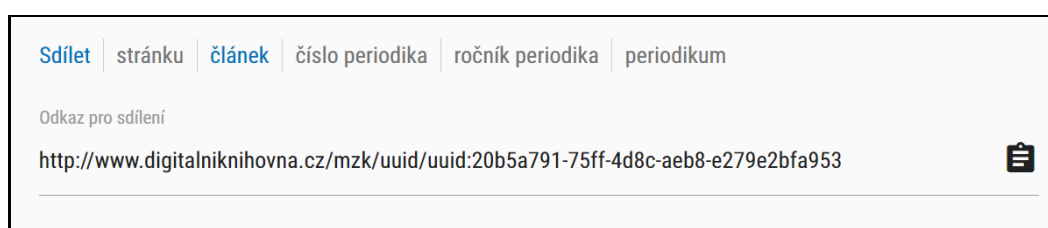
⁸ BELLINI, Emanuele, Chiara CIRINNA a Maurizio LUNGHI. Trvalé identifikátory pro kulturní dědictví. *Digital Preservation Europe* [online]. 2008 [cit. 2021-10-01]. Dostupné z: https://issuu.com/dpe-staff/docs/cz_trvale_identifikatory

UUID se skládá z 36 znaků, respektive z 32 alfanumerických znaků, které jsou oddělené 4 spojovníky. Pro 32 alfanumerických znaků se používá [šestnáctková soustava](#) (0–9, A–F) a znaky jsou pomocí spojovníků rozděleny do 5 částí ve tvaru 8-4-4-4-12.

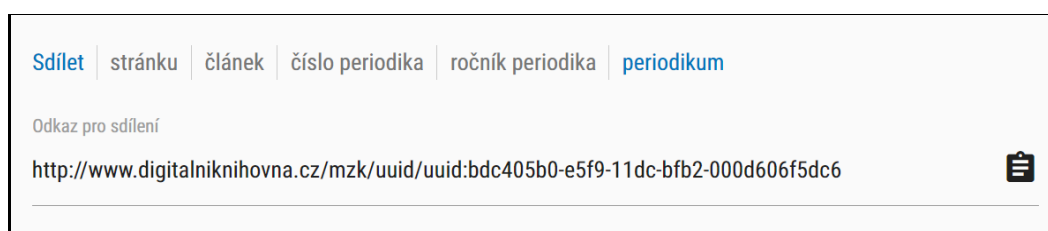
Např. 18af73d0-5ade-11e6-8e4e-005056827e51

Více informací ke generování UUID se nachází [zde](#).

UUID se využívá například v systému Kramerius, což je systém pro zpřístupňování digitálních dokumentů, který se používá v mnoha českých knihovnách a institucích. Pomocí UUID se dají identifikovat jak celé tituly, tak i ročníky, čísla, články, strany či virtuální sbírky. Identifikátor UUID se pak zobrazuje v URL odkazu, pomocí kterého se lze dostat na daný digitální objekt (viz obrázky níže).



Obrázek 17

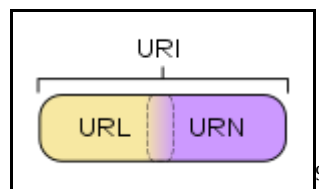


Obrázek 18

1.3.2 URI

URI (Uniform Resource Identifier = jednotný identifikátor zdroje) je **soubor znaků, který slouží k přesné identifikaci zdroje informací v prostředí počítačových sítí**. Těmito zdroji mohou být dokumenty, knihy, obrázky, databáze, softwary, osoby, ale i koncepty.

URI může popisovat zdroj jak čistě z hlediska jeho identity (a nepopisuje, jak se ke zdroji dostat), tak čistě z hlediska lokace zdroje (a pak nepopisuje identitu zdroje), tak z obou hledisek najednou (je popsána přesná identita zdroje a jeho lokace). Identita zdroje se popisuje pomocí identifikátoru URN a lokace zdroje se popisuje pomocí identifikátoru URL. Oba tyto identifikátory si blíže popíšeme dále ve studijním textu.



Obrázek 19

URI je velmi obecný koncept, a proto je jeho základní formát velmi volný. Každé URI začíná názvem schématu a dvojtečkou. Poté následuje řetězec znaků, jehož význam a formátování závisí na použitém schématu (jinými slovy, použité schéma odkazuje na specifikaci pro vytváření další části URI). Každé schéma musí být registrované v [registru URI schémat](#), o který se stará instituce IANA (Internet Assigned Numbers Authority). Jak je z registru vidět, URI schémat existuje velké množství. Některá schémata jsou zároveň i protokoly (o protokolech si řekneme více u identifikátoru URL), například HTTP, HTTPS či FTP. Své schéma má například i DOI (viz kapitola DOI) či služba Spotify.

Příklady URI:

`https://www.knihovny.cz/`

`doi:10.1108/02621719710174525`

`spotify:track:4fqhx5QGIZXQO3xLmUNZOM`

1.3.3 URL

URL (Uniform Resource Locator = jednotný lokátor zdroje) je **soubor znaků, který slouží k identifikaci přesného umístění informací na Internetu**. V internetovém prohlížeči se zadává do adresního řádku.

Jak již víme, URL identifikuje pouze umístění a ne zdroj samotný. Nevýhodou URL je to, že **když se změní umístění zdroje, URL již není platné a cesta ke zdroji je ztracena**.

URL se skládá z následujících částí:

- **protokol** – Protokoly určují způsob přenosu dat. Nejčastěji se setkáte s protokoly HTTP a HTTPS (zapisují se `http://` a `https://`). Tyto dva protokoly se používají pro HTML stránky. Liší se od sebe tím, že HTTPS je protokol šifrovaný (S = security). Poté existují i další protokoly jako `mailto:` a `ftp://`, ale ty se užívají zřídka.
- **doména** – Je to jméno, pod kterým lze nalézt webové stránky. Doménové jméno se skládá ze tří částí oddělených tečkami:

⁹ Obrázek převzat z Uniform Resource Identifier. *Wikipedie* [online]. [cit. 2021-9-4]. Dostupné z: https://cs.wikipedia.org/wiki/Uniform_Resource_Identifier

- **doména třetího řádu** – Jméno virtuálního serveru, většinou to je „www“. Správně nastavený web by měl fungovat i bez zadání „www“. Někdy se můžete setkat s případy, kdy je místo „www“ něco jiného.
- **doména druhého řádu** – Je třeba ji registrovat. Mezi nejznámější patří google, seznam, heureka atd.
- **doména nejvyššího řádu** – Vyjadřuje příslušnost k zemi (cz, sk, de) nebo určuje zaměření (com, gov, org, edu). Doménu cz spravuje sdružení [NIC](#).
- **port** – Velmi zřídka se za doménu píše dvojtečka a číslo portu. Používá se to hlavně u vývojových verzí nebo na intranetech. Pokud jméno portu uvedeno není, server odpovídá na portu 80 pro protokol HTTP a na portu 443 pro protokol HTTPS.
- **adresářová cesta** – Soubory jsou často uloženy v adresáři. Cesta do adresáře začíná vždy lomítkem, poté následuje název adresáře a pak zase lomítko. Je-li adresářů více úrovní, píšou se za sebe a jsou odděleny lomítky.
- **soubory** – Jméno souboru se píše za adresářovou cestu.
- **dotaz** – Součástí URL mohou být i vstupní data pro nějaký skript. Ta se do URL píšou za otazník. Syntax je ?jmeno=hodnota&jmeno2=hodnota2.
- **záložka** – Neboli fragment či kotva. Pomocí křížku # můžeme odkazovat na konkrétní místo v dokumentu, tedy na záložku. Křížek se píše za název souboru a za křížek se píše název záložky.

Pokud by se v URL nacházely všechny části, syntax by vypadala takto:

protokol://doména3.doména2.doména1:port/cesta/soubor?dotaz#záložka

Příklady URL:

https://cs.wikipedia.org/wiki/Uniform_Resource_Locator#Extern%C3%AD_odkazy

<https://www.digitalniknihovna.cz/mzk/about>

<https://www.nkp.cz/digitalni-knihovna/dnnt/NDK-DNNT-FAQ#1>

Výše popsaným adresám se říká absolutní. Pak se ještě používají relativní adresy, které vychází z umístění dokumentu. U relativních adres můžeme některé části adresy vynechat, např. když je stránka umístěna na stejném serveru, můžeme vynechat jméno serveru. Jestliže je umístěna ve stejném adresáři, není potřeba psát ani cestu.

1.3.4 PURL

PURL (Persistent Uniform Resource Locator) je **trvalý identifikátor** vyvinutý institucí OCLC, který na rozdíl od URL **nevede přímo na konkrétní stránku, ale na resolver** (přesměrovávací službu). V resolveru jsou obsaženy registry, v nichž jsou jednotlivé PURL provázané s aktuálními URL, takže uživatel je skrze PURL přesměrován na aktuální URL. **Pokud zdroj změní lokaci a změní se URL, uživatel se přes PURL**

ke zdroji i nadále dostane. Pro přístup k PURL resolverům a směrování URL odkazů zpět do klientského prohlížeče je využíváno standardního HTTP protokolu.

Instituce mohou využívat resolver provozovaný institucí Internet Archive na adrese <https://archive.org/services/purl/> (dříve byl resolver provozovaný institucí OCLC), nebo si mohou spravovat své vlastní resolvers. V každé instituci by měl být správce PURL identifikátorů, který v resolveru generuje PURL a udržuje aktuálnost dat.

PURL se skládá z 3 částí, a to z protokolu, adresy PURL resolveru a jména/identifikátoru PURL.

Např. <http://purl.org/dc/terms>

- protokol: http://
- adresa PURL resolveru: purl.org
- jméno/identifikátor PURL: /dc/terms

1.3.5 URN

URN (Uniform Resource Name = jednotné jméno zdroje) slouží jako **trvalý identifikátor, který identifikuje zdroj nezávisle na jeho umístění**. Jakmile je jednou URN danému zdroji přiřazeno, již se nikdy nemění. Pomocí resolveru můžeme přeložit URN na URL a získáme tak lokaci zdroje.

Každé URN začíná prefixem „urn:“, poté následuje jmenný prostor identifikátoru (NID – Namespace Identifier) zakončený dvojtečkou a nakonec se uvádí specifický řetězec jmenného prostoru (NSS – Namespace Specific String).

urn:NID:NSS

Jmenný prostor identifikátoru může být tvořen písmeny (nerozlišují se velká a malá), čísly a spojovníky. Každý jmenný prostor identifikátoru musí být registrován v [registru URN jmenných prostorů](#), o který se stará instituce IANA (Internet Assigned Numbers Authority). Mezi registrované jmenné prostory patří například ISBN, ISSN či NBN, o kterém si více řekneme v kapitole URN:NBN. Podoba specifického řetězce jmenného prostoru je určena použitým jmenným prostorem identifikátoru (instituce, která se stará o jmenný prostor identifikátoru, si určí, jaký má mít specifický řetězec jmenného prostoru syntax).

Příklady URN:

urn:isbn:0-486-27557-4

urn:issn:0167-6423

urn:oasis:names:specification:docbook:dtd:xml:4.1.2

1.3.6 URN:NBN

Jak jsme si řekli v minulé kapitole, jedním z identifikátorů jmenného prostoru u URN je URN:NBN. Jedná se o globální systém, který slouží pro trvalou identifikaci digitálních dokumentů a který též podporuje zpřístupnění a dlouhodobé uchovávání digitálních dokumentů. Původně byl identifikátor URN:NBN zamýšlen jen pro digitální dokumenty spadající do národní bibliografie, později se jeho užití rozšířilo i na další dokumenty z oblasti digitálního kulturního dědictví.¹⁰

Systém URN:NBN mohou v jednotlivých zemích zavádět pouze národní knihovny, které zároveň vytváří pravidla pro používání v dané zemi. V ČR je centrální autoritou tohoto systému Národní knihovna ČR, která provozuje nástroj [CZIDLO](#) (Czech IDentification and LOcalization tool = Český systém pro identifikaci a lokalizaci dokumentů digitálního kulturního dědictví). Tento nástroj slouží pro přidělování identifikátorů URN:NBN (automatizovaně prostřednictvím API nebo manuálně přes webové rozhraní), správu identifikátorů (např. deaktivování), dodatečné vkládání adres URL digitálních dokumentů a také slouží jako resolver zajišťující přesměrovávání webového prohlížeče z URN:NBN na aktuální URL umístění dokumentu. CZIDLO tak zajišťuje trvalý přístup k digitálním dokumentům navzdory změnám internetových adres.

URN:NBN může být přiděleno takovému digitálnímu dokumentu, který lze považovat za součást českého kulturního dědictví a který bude dlouhodobě archivován a trvale zpřístupňován v digitální knihovně na nekomerční bázi. Digitální dokument musí mít neměnnou povahu z hlediska obsahu (např. aktualizované články jsou z identifikace vyloučeny). Podrobnější pravidla pro přidělování URN:NBN se nachází na [webu nástroje CZIDLO](#) pod záložkou „Pravidla“.

Syntax identifikátoru je taková, že na začátku se nachází označení „urn:nbn“, dále následuje kód země (pro ČR se používá „cz“), poté se uvádí kód vlastníka dokumentu a poslední částí je šestimístný alfanumerický řetězec, který je generován resolverem. Z URN:NBN lze tedy zjistit vlastníka dokumentu. Vlastník dokumentu bývá označován pomocí sigly knihovny; Národní knihovna ČR, Moravská zemská knihovna v Brně a Vědecká knihovna v Olomouci používají místo sigly vlastní označení (nk, mzk, vko). Všechny části jsou oddělené dvojtečkou, pouze šestimístný řetězec je oddělen spojovníkem.

Příklady URN:NBN identifikátorů:

urn:nbn:cz:mzk-005i1f

urn:nbn:cz:aba013-0003f9

urn:nbn:cz:vko1-00a9ok

¹⁰ CUBR, Ladislav a Zdeněk VAŠEK. Identifikátory digitálních dokumentů se zaměřením na systém URN:NBN v ČR. *Čtenář: Měsíčník pro knihovny* [online]. 2013, 65(6) [cit. 2021-10-03]. Dostupné z: <https://svvkl.cz/en/ctenar/clanek/1254>

V nástroji [CZIDLO](#) si můžete zkusit dané identifikátory vyhledat. Naleznete tam jak informace o daném fyzickém dokumentu, tak technické informace o digitálním dokumentu, vlastníka digitálního dokumentu a URL odkaz na digitální dokument.

1.3.7 HANDLE

Systém Handle byl vyvinut Korporací pro národní výzkumné iniciativy (CNRI – Corporation for National Research Initiative) a jedná se o **identifikační mechanismus pro přiřazování, správu a řešení trvalých identifikátorů pro digitální objekty a jiné internetové zdroje**. Handle systém využívá směrovací server, který zajišťuje, že **uživatel bude vždy odkázán na aktuální URL odkaz digitálního objektu** (vlastník dokumentu musí URL aktualizovat). Identifikátory Handle tak nejsou založené na proměnlivých vlastnostech zdroje, jako je lokace či vlastnictví. Pokud dojde ke změně lokace, identifikátor Handle zůstává stejný.

Architektura systému Handle má dvoustupňový hierarchický model služeb. První úroveň je globální registr Handle (GHR – Global Handle Registry) a druhá úroveň je tvořena všemi dalšími místními službami Handle (LHS – Local Handle Services). Globální registr se nazývá HANDLE.NET a slouží k registraci a správě prefixů Handle identifikátorů. Tento registr dále obsahuje informace o všech LHS. LHS jsou provozované institucemi s administrativní zodpovědností za tvorbu a údržbu identifikátorů pod určitým prefixem. Když uživatel odešle dotaz na digitální objekt identifikovaný identifikátorem Handle, globální registr ho přesměruje na konkrétní LHS, který tento identifikátor spravuje, a LHS odešle záznam daného identifikátoru do prohlížeče uživatele.¹¹

Identifikátor Handle se skládá z prefixu a sufixu (jsou oddělené lomítkem). Prefix je globálně jedinečný řetězec čísel, který identifikuje instituci přidávající identifikátory. Jak již víme, prefix je přidělován globálním registrem Handle. Sufix identifikuje konkrétní informační zdroj. Každý sufix musí být jedinečný v rámci svého prefixu — tím je zajištěno, že každý takto tvořený identifikátor Handle bude jedinečný.

Příklady Handle identifikátorů:

20.1000/100

4263537/5555

10.1000/182

Handle identifikátor může být vyjádřen v podobě URL:

<http://hdl.handle.net/4263537/5555>

Směrování Handle identifikátorů si můžete vyzkoušet na [této stránce](#) (můžete tam zkusit vložit jeden z výše uvedených identifikátorů).

¹¹ JÍLKOVÁ, Marta. *Persistentní identifikátory a jejich využití v digitálních knihovnách* [online]. Praha, 2013, s. 28-29 [cit. 2021-9-4]. Dostupné z: <https://dspace.cuni.cz/bitstream/handle/20.500.11956/51726/120138608.pdf?sequence=1&isAllowed=y>. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta. Vedoucí práce Barbora Drobíková.

Systém Handle je komerční, platí se za registraci prefixů a za údržbu prefixů v globálním registru Handle.

1.3.8 DOI

DOI (Digital Object Identifier = identifikátor digitálního objektu) je **trvalý identifikátor digitálního objektu v elektronickém prostředí**. Systém DOI je vystavěn na základě systému Handle. DOI tedy jednoznačně identifikuje digitální objekt a zároveň k němu poskytuje trvalý odkaz. Jednou přidělený identifikátor DOI je trvalý; když se digitální objekt přesune na jinou lokaci na webu, identifikátor zůstává stejný a uživatel se k digitálnímu objektu i nadále dostane.

DOI je **využíván zejména vydavateli vědeckých publikací pro identifikaci jejich článků**. Systém DOI je komerční; instituce, které přidělují svým publikacím DOI, platí roční poplatek a poté platí za každý přidělený identifikátor dané registrační agentuře (např. Crossref, jejich aktuální ceník se nachází [zde](#)). Pro koncového uživatele, který DOI používá pro vyhledávání, je služba bezplatná.

DOI je jedno z registrovaných URI schémat. Už víme, že každé URI začíná názvem schématu a dvojtečkou; před samotný identifikátor DOI se tedy píše „doi:“. Poté následuje identifikátor DOI, který se skládá z prefixu a sufixu (ty jsou oddělené lomítkem). Prefix je údaj, který jednoznačně identifikuje instituci, která daný digitální objekt opatřila číslem DOI (prefix přiděluje registrační agentura). Prefix musí vždy začínat číslem 10, protože tím se DOI odlišuje od ostatních implementací systému Handle. K přidělenému prefixu si instituce sama určí svůj vlastní sufix, což je libovolný řetězec znaků, který jednoznačně identifikuje digitální objekt.

Příklady DOI identifikátorů:

doi:10.2824/36822

doi:10.1002/0470841559.ch1

doi:10.5817/CZ.MUNI.P210-9822-2020

DOI může být vyjádřeno jako URL prostřednictvím protokolu HTTP. Pak se z DOI stává proklikatelný, interoperabilní a trvalý odkaz (tedy PURL). Např. <https://doi.org/10.2824/36822>

Majitel digitálního objektu vkládá přidělené DOI do databáze registrační agentury a doplňuje metadata o digitálním objektu (včetně URL odkazu na aktuální lokaci digitálního objektu). Jestliže se digitální objekt přesune na jinou lokaci, je jeho vlastník povinen aktualizovat URL digitálního objektu v databázi registrační agentury. Registrační agentura pak zajišťuje směřování DOI identifikátorů: tj. po obdržení DOI identifikátoru vrátí uživateli webovou stránku, která obsahuje plný text digitálního objektu nebo (není-li plný text volně dostupný) jeho bibliografickou citaci a informaci o tom, jak lze plný text objektu získat. Směřování DOI identifikátorů si můžete vyzkoušet na [této stránce](#).

1.3.9 ARK

ARK (Archival Resource Key = klíč archivního zdroje) je **trvalý identifikátor určený k podpoře dlouhodobého přístupu k informačním objektům**. ARK se používá pro **digitální objekty** (např. digitální dokumenty, databáze, obrázky, software, webové stránky), **fyzické předměty** (např. knihy, sochy), **živé bytosti a skupiny** (např. lidé, zvířata, firmy, orchestry) a **nehmotné předměty** (např. nemoci, místa, chemikálie, slovní zásoby).¹² Využívají ho především instituce zabývající se ochranou digitálních dat, knihovny, archivy a další paměťové instituce. Výhodou ARK identifikátoru oproti předchozím identifikátorům je to, že **za přidělování a užívání ARK se neplatí žádné poplatky**.

Identifikátor ARK funguje jako specifické URL zajišťující trvalý a dlouhodobý přístup k informačnímu zdroji. Identifikátor se váže k danému informačnímu objektu, nikoliv k umístění, a proto ho lze snadno přesunout na nové servery bez ztráty funkčnosti, přičemž vlastní trvalá část identifikátoru zůstává úplně stejná. V URL se nejprve nachází protokol HTTP nebo HTTPS a poté jmenný prostor konkrétní instituce, která poskytuje podporu pro daný ARK (NMA = Name Mapping Authority). Tato část je volitelná a měnitelná. Poté následuje neměnný identifikátor ARK, který začíná označením „arc:“. Následně se uvádí pětimístné číslo instituce, která identifikátory přiřazuje (**NAAN** = Name Assigning Authority Number). Pro získání čísla instituce (NAAN) se musí žadatelé [zaregistrovat v Kalifornské digitální knihovně](#), která čísla přiděluje a která se stará o registr NAAN. Poté se uvádí jméno ARK, tedy automaticky generovaný řetězec znaků, který identifikuje samotný informační objekt (ARK name). Pro generování jména ARK lze využít různé open source softwary (příkladem je třeba služba [EZID](#), kterou provozuje Kalifornská digitální knihovna, nebo služba [NOID](#)). Na konec identifikátoru lze pak ještě vložit kvalifikátor (Qualifier), který slouží k odlišení různých verzí či formátů informačního objektu (např. TIFF a JPEG), či k rozlišení částí informačního objektu (titulní listy, konkrétní stránky, kapitoly).¹³

Pokud by se v ARK nacházely všechny části, syntax by vypadala takto:

<http://NMA/ark:/NAAN/Name/Qualifier>

Příklady ARK:

<http://example.org/ark:/12025/654xz321/s3/f8.05v.tiff>

<http://gallica.bnf.fr/ark:/12148/bpt6k94258n.image.f126.tableDesMatières>

¹² ARK. *Confluence: CDL Curation* [online]. 22. 3. 2018 [cit. 2021-10-5]. Dostupné z: <https://confluence.ucop.edu/display/Curation/ARK>

¹³ JÍLKOVÁ, Marta. *Persistentní identifikátory a jejich využití v digitálních knihovnách* [online]. Praha, 2013, s. 38-39 [cit. 2021-9-4]. Dostupné z: <https://dspace.cuni.cz/bitstream/handle/20.500.11956/51726/120138608.pdf?sequence=1&isAllowed=y>. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta. Vedoucí práce Barbora Drobíková.

2 METADATA

Stručně řečeno, metadata jsou data o datech. Metadata z pohledu knihovnictví by se dala definovat takto: „**Metadata jsou strukturované informace, které popisují, vysvětlují, lokalizují nebo jinak usnadňují vyhledávání, používání nebo správu informačního zdroje.**“¹⁴ Typickým příkladem metadat v knihovnictví je katalogizační záznam, který nám nejen popisuje zdroj, ale zároveň nám říká i jeho lokaci. Uživatel nemusí procházet všechny regály v knihovně, aby našel požadované dokumenty, ale vyhledává pouze v knihovním katalogu — tedy v metadatach. Pomocí metadat můžeme popisovat jak fyzické (analogové) zdroje, tak i digitální zdroje.

V rámci digitalizace dokument naskenujeme a skeny poté opatříme metadaty. Skeny a metadata pak dohromady tvoří digitální objekt, který můžeme zpřístupňovat, ukládat v dlouhodobém úložišti anebo vyměňovat s jinými institucemi. Pomocí metadat můžeme vyjádřit různé vlastnosti digitálního objektu. Podle těchto vlastností můžeme metadata dělit dle těchto typů:

- **popisná metadata** – popisují digitální objekt a umožňují jeho vyhledání,
- **administrativní metadata** – poskytují informace, které pomáhají spravovat digitální objekt; řadí se sem technická metadata, archivační metadata a právní metadata,
- **strukturální metadata** – popisují vnitřní strukturu digitálního objektu.

Podle typů metadat vznikají jednotlivé metadatové formáty, na základě kterých se pak vytváří metadatové záznamy s daným typem metadat. V dalších kapitolách se na jednotlivé typy metadat podíváme blíže a u každého typu si představíme formáty, které jsou považovány za mezinárodní standardy v oblasti digitalizace a digitální archivace. Tyto formáty jsou využívány i v projektu Národní digitální knihovna a v dalších českých digitalizačních projektech. Ještě předtím si ale vysvětlíme, co to je metadatové schéma a XML.

2.1 Metadatové schéma

Metadatové schéma je „**souhrn metadatových elementů, jejich struktury a vlastností, které jsou určeny k popisu primárních dat v určité specifické oblasti.**“¹⁵ Metadatová schémata vytváříme a používáme proto, abychom byli při popisu primárních dat jednotní. Existují různá standardizovaná mezinárodní schémata zastřešená významnými organizacemi. Dodržování standardizovaných schémat umožňuje snadnou výměnu dat mezi institucemi. I my sami si však můžeme vytvořit vlastní metadatové schéma.

¹⁴ NISO. *Understanding Metadata* [online]. 2004, s. 1 [cit. 2021-9-27]. ISBN 1-880124-62-9. Dostupné z: https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf

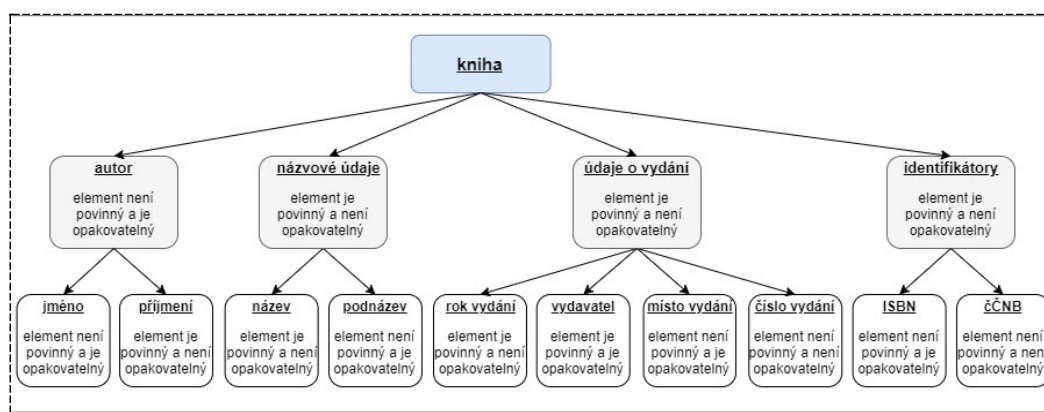
¹⁵ RYCHTÁŘOVÁ, Pavla, KRAVECOVÁ, Lenka a Markéta KRUTSKÁ: *Metadata I*. E-learningový kurz Digitalizace v knihovnách [online]. Brno: Moravská zemská knihovna v Brně, 2021 [cit. 2021-8-22]. Dostupné z: <https://kurzy.knihovna.cz/>

Nyní si zkusíme vytvořit jednoduché schéma, pomocí kterého bychom si mohli popsat knihy v naší osobní knihovně. Nejprve si musíme zvolit elementy, pomocí kterých si knihy popíšeme. Elementy budou jméno, příjmení, název, podnázev, rok vydání, vydavatel, místo vydání, číslo vydání, ISBN a čČNB. Elementy si můžeme vyjádřit pomocí této struktury:

```

kniha
  autor
    jméno
    příjmení
  názvové údaje
    název
    podnázev
  údaje o vydání
    rok vydání
    vydavatel
    místo vydání
    číslo vydání
  identifikátory
    ISBN
    čČNB
  
```

Dále můžeme každý element definovat vlastnostmi povinnost a opakovatelnost. Povinnost nám říká, zda element musíme použít vždy, či zda může být v některých případech vynechán (starší knihy například nemají ISBN, proto element není povinný). Opakovatelnost nám říká, zda můžeme nějaký element použít vícekrát (některé knihy mají více autorů, proto je tento element opakovatelný). Na obrázku níže je graficky znázorněna struktura schématu a u každého elementu je definována povinnost a opakovatelnost.



Obrázek 20 – Metadatové schéma

Nyní už víme, kam co zapsat. Dále si stanovíme pravidla popisu, která nám budou říkat, jakým způsobem máme metadata zapsat. Díky tomu předejdeme situaci, že by se stejná informace zapisovala různými způsoby (například místo vydání by jednou bylo „Brno“ a podruhé „V Brně“).

Jméno: Pokud má autor dvě či více křestních jmen, zapíšeme je za sebou v celém znění.
Příjmení: Pokud nedokážeme rozlišit jméno a příjmení (např. Marie Terezie, Jan z Arrasu), vše zapisujeme do příjmení.
Název: Zapisujeme tak, jak je uveden na titulním listu.
Podnázev: Zapisujeme tak, jak je uveden na titulním listu.
Rok vydání: Zapisujeme ve formátu RRRR. Pokud rok vydání není uveden, zapíšeme alespoň desetiletí (pokud víme) anebo století ve formátu [RRR-], [RR--].
Vydavatel: Pokud je vydavatel uveden pomocí zkratky, tak do metadat ho zapíšeme v celém znění.
Místo vydání: Zapisujeme v 1. pádu, přestože je třeba v knize napsané „V Brně“.
Číslo vydání: Zapisujeme vždy pomocí číslovky, přestože je třeba v knize napsané „vydání sedmé“.
ISBN: Zapisujeme bez slova „ISBN“.
čČNB: Zjišťujeme v bázi ČNB a zapisujeme i s prefixem „cnb“.

Obrázek 21 – Pravidla popisu

Nyní nám již nic nebrání v tom, abychom si pomocí schématu popsali jednu konkrétní knihu. Níže můžeme vidět příklad. Všimněte si, že byla dodržena struktura i pravidla, a že byly použity všechny povinné elementy. V tomto konkrétním případě nebyl použit nepovinný element podnázev. Jeden element se nám tam objevil dvakrát, a to autor, což je v pořádku, protože tento element je opakovatelný.

kniha:		
autor:	jméno:	Hugo
	příjmení:	Mercier
autor:	jméno:	Dan
	příjmení:	Sperber
názvové údaje:	název:	Záhada rozumu
údaje o vydání:	rok vydání:	2019
	vydavatel:	Host
	místo vydání:	Brno
	číslo vydání:	1
identifikátor:	ISBN:	978-80-7577-635-8
	čČNB:	cnb003115975

Obrázek 22 – Příklad

V rámci digitalizace často potřebujeme metadatová schémata zpracovávat strojově. Je tedy dobré mít schémata v podobě, se kterou si stroje dokáží poradit. Tou podobou je například značkovací jazyk XML.

2.2 XML

Mnoho metadatových schémat, která se používají v rámci digitalizace, je založených na jazyku XML, proto je dobré znát základy a syntaxi tohoto jazyka. XML (Extensible Markup Language = rozšiřitelný značkovací jazyk) je **obecný značkovací jazyk**, který byl vyvinut konsorciem W3C. Je **určen především pro výměnu dat mezi aplikacemi**

a pro publikování dokumentů. XML popisuje strukturu z hlediska věcného obsahu jednotlivých částí, nezabývá se vzhledem (vzhled pak může být definován pomocí [kaskádových stylů = CSS](#)). XML je otevřený formát a je podporován řadou nástrojů a programovacích jazyků.

XML se dělí do částí, kterým se říká elementy. Elementy do sebe mohou být vnořené a tím se může vyjádřit struktura informací v dokumentu (např. při popisu knihy bychom měli elementy pro kapitoly a každá kapitola by pak obsahovala element nadpis a několik elementů pro odstavce). XML musí mít právě jeden rodičovský element (kořenový/root element), do kterého mohou být vnořeny další elementy. Elementy se ale nemohou překrývat; to znamená, že každý element musí být celý obsažen v jiném elementu (kromě kořenového elementu). Každý element má počáteční tag (<kniha>) a ukončovací tag (</kniha>). Pomocí tagů popisujeme význam jednotlivých částí. Pozor, u tagů se rozlišují velká a malá písmena. U tagů se též nepoužívá diakritika (v obsahu elementů je samozřejmě diakritika v pořádku). XML neobsahuje předdefinované tagy, je třeba si definovat vlastní tagy, které budeme používat, anebo můžeme použít existující schémata, kde jsou tagy definovány.

```
<kniha>
  <kapitola>
    <nadpis>Zde je název kapitoly.</nadpis>
    <autor>Jan Novák</autor>
    <odstavec>Zde je první odstavec.</odstavec>
    <odstavec>Zde je druhý odstavec.</odstavec>
  </kapitola>
</kniha>
```

Některé elementy nemusí mít žádný obsah. Takový element můžeme popsat tak, že uvedeme počáteční tag a pak hned ukončovací.

```
<nadpis></nadpis>
```

To ale není úplně pohodlné, a proto se většinou prázdné elementy označují tagem „prázdný element“. V takovém případě se uvede pouze jeden tag, který má na konci názvu lomítko.

```
<nadpis/>
```

U každého počátečního tagu pak ještě můžeme použít atributy. Atributy se používají k upřesnění významu elementu, tedy k přidání dalších důležitých informací. Hodnotu atributu musíme vždy uzavřít do jednoduchých nebo dvojíтых uvozovek (ale jednoduchá uvozovka musí být uzavřena jednoduchou a dvojíтая dvojíтой). U jednoho tagu lze použít více atributů (oddělíme je mezerou).

```
<kapitola autor="Jan Novák">
  ....
</kapitola>
```

Do prvního řádku XML dokumentu se ještě uvádí deklarace verze XML (atribut version) a znakové sady (atribut encoding). Co se týká verzí XML, tak aktuální verze XML je 1.1 a pak ještě existuje první verze 1.0 v páté revizi (více o rozdílu mezi verzemi [zde](#)). Aby šlo v XML bez problémů zapisovat znaky libovolného jazyka,

je potřeba definovat znakovou sadu. V českém prostředí se obvykle používá kódování [UTF-8](#).

Níže je příklad XML dokumentu včetně popisu terminologie. Bylo použito stejné metadatové schéma, které se nachází i v předchozí kapitole.

<code><?xml version="1.0" encoding="UTF-8"?></code>	→ deklarace verze XML a znakové sady
<code><kniha></code>	→ root/kořenový element
<code><autor typ="osoba"></code>	→ název atributu a hodnota atributu
<code><jmeno>Hugo</jmeno></code>	→ element
<code><prijmeni>Mercier</prijmeni></code>	→ počáteční tag a ukončovací tag
<code></autor></code>	
<code><autor typ="osoba"></code>	
<code><jmeno>Dan</jmeno></code>	
<code><prijmeni>Sperber</prijmeni></code>	
<code></autor></code>	
<code><nazvove_udaje></code>	
<code><nazev>Záhada rozumu</nazev></code>	→ obsah elementu
<code><podnazev/></code>	→ prázdný element
<code></nazvove_udaje></code>	
<code><udaje_o_vydani></code>	
<code><rok_vydani>2019</rok_vydani></code>	
<code><vydavatel>Host</vydavatel></code>	
<code><misto_vydani>Brno</misto_vydani></code>	
<code><cislo_vydani>1</cislo_vydani></code>	
<code></udaje_o_vydani></code>	
<code><identifikatory></code>	
<code><ISBN> 978-80-7577-635-8</ISBN></code>	
<code><cCNB>cnb003115975</cCNB></code>	
<code></identifikatory></code>	
<code></kniha></code>	

Obrázek 23 – Příklad XML dokumentu

Již víme, že při vytváření XML dokumentů musíme dodržovat určitá pravidla. Níže jsou uvedené 4 ukázky XML, ve kterých se nachází jedna či dvě chyby. Zkuste si sami určit, kde se chyby nachází, poté se můžete podívat na řešení.

`<autor><jmeno>Anna</jmeno><prijmeni>Veselá<autor>`

→ Zde jsou 2 chyby. U elementu *prijmeni* chybí ukončovací tag a u ukončovacího tagu *autor* chybí lomítko.

`<kniha><kapitola>Obsah kapitoly</kniha></kapitola>`

→ V tomto případě je chyba v tom, že elementy jsou překříženy. Ukončovací tagy by měly být v opačném pořadí, nejprve by se měl ukončit tag *kapitola* a poté tag *kniha* (protože element *kapitola* je vnořen do elementu *kniha*).

`<rok_vydani>2005</rok_vydani><misto_vydani>Praha</misto_vydani>`

→ Zde chybí kořenový element. XML dokument musí mít právě jeden kořenový element, který „obalí“ všechny ostatní elementy.

`<Autor typ="žena"><jmeno>Barbora</jmeno><prijmeni>Tichá</prijmeni></autor>`

→ V tomto případě je chybně zapsaná hodnota atributu, jednoduchá uvozovka musí být uzavřena jednoduchou a dvojitou. Dále je zde chyba v tom, že první tag začíná velkým písmenem a ukončovací tag začíná malým písmenem, takže se ve skutečnosti jedná o dva různé tagy.

Validátory XML

Na Internetu existují různé XML validátory, pomocí kterých si můžete zkontrolovat, zda máte v XML dokumentu správný syntax. Můžete si vyzkoušet například [tento](#) anebo [tento](#) validátor, ale existuje jich mnohem více.

2.2.1 XSD

XSD (XML Schema) je **jazyk, který slouží pro popis struktury dokumentu XML**. XSD definuje:

- elementy a atributy, které se mohou objevit v XML,
- které elementy jsou „podelementy“ jiných,
- pořadí a počet „podelementů“,
- „obsah“ elementu (zda je prázdný či obsahuje text),
- datový typ elementů a atributů,
- výchozí a fixní hodnoty elementů a atributů.

Abychom mohli XML dokument označit za validní, musí odpovídat svému XSD. V dalších kapitolách tohoto studijního textu se budeme zabývat konkrétními metadatovými schémata. Tato schémata mají definované své XSD. Při práci s konkrétním metadatovým schématem musíme respektovat dané XSD a například si nemůžeme vymýšlet nové elementy či atributy, které nejsou definované v XSD. Alternativou XSD je jazyk DTD (Document Type Definition), který se ale dnes již tolik nepoužívá.

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="kniha">
    <xs:complexType>
      <xs:sequence>
        <xs:element maxOccurs="unbounded" name="autor">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="jmeno" type="xs:string" />
              <xs:element name="prijmeni" type="xs:string" />
            </xs:sequence>
            <xs:attribute name="typ" type="xs:string" use="required" />
          </xs:complexType>
        </xs:element>
        <xs:element name="nazvove_udaje">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="nazev" type="xs:string" />
              <xs:element name="podnazev" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```

    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="udaje_o_vydani">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="rok_vydani" type="xs:unsignedShort" />
      <xs:element name="vydavatel" type="xs:string" />
      <xs:element name="misto_vydani" type="xs:string" />
      <xs:element name="cislo_vydani" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="identifikatory">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="ISBN" type="xs:string" />
      <xs:element name="cCNB" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>

```

Obrázek 24 – XSD pro XML dokument, který jsme si popisovali dříve

Již víme, co je to metadatové schéma, a rozumíme značkovacímu jazyku XML. Nyní se můžeme podívat na konkrétní příklady metadatových formátů.

2.3 Popisná metadata

Popisná metadata **popisují a identifikují daný objekt, a tím usnadňují jeho vyhledání**. Popisná metadata u digitálních knihovních dokumentů zahrnují jmenný popis (název, autor, nakladatelské údaje), věcný popis (klíčová slova, předmětová hesla), popis fyzického dokumentu (počet stran, rozměry) a identifikátory dokumentu (ISBN/ISSN/ISMN, čárový kód, signatura, čČNB, UUID, URN:NBN).

Výše uvedené údaje bývají většinou zapsané v katalogizačním záznamu. Při digitalizaci se proto tyto údaje přebírají z katalogu a není potřeba je vytvářet znovu. Postup je takový, že z katalogu se stáhne katalogizační záznam a zkonvertuje se do potřebného formátu. Výjimku tvoří identifikátory UUID a URN:NBN, které se vytváří až během digitalizace.

2.3.1 MARC 21, MARCXML

MARC 21 je mezinárodní formát, který se **používá pro katalogizaci knihovních dokumentů**. Tento formát využívá třímístná numerická pole a jejich podpole pro podrobný popis dokumentů. V českém prostředí je tento formát velmi používaný a jak již víme, popisná metadata u digitálních dokumentů se často přebírají právě z katalogizačních záznamů. Je proto užitečné znát strukturu tohoto formátu.

O formátu MARC 21 se můžete více dočíst na stránkách [Kongresové knihovny](#), která tento formát spravuje, anebo na stránkách [Národní knihovny ČR](#). Zároveň doporučujeme k prostudování i pravidla popisu [RDA](#) (dříve se používala pravidla AACR2).

Aby se mohl formát MARC 21 konvertovat do jiných formátů, vznikl formát [MARCXML](#), který zachovává všechna pole MARC 21 a prezentuje je prostřednictvím jazyka XML.

LDR	-----nam-a22-----a-4500
001	001104126
003	CZ-BrMZK
005	20170525072947.0
008	090903s2009----xr-a---g-----000-f-cze--
015	a cnb001997242
020	a 978-80-87209-32-5 q (váz.)
020	a 978-80-87209-33-2 q (váz. + kalendář)
040	a BOA001 b cze
072 7	a 821.162.3-3 x Česká próza 2 Konspekt 9 25
080	a 821.162.3-3 2 MRF
1001	a Čapek, Karel, d 1890-1938 7 jk01021023 4 aut
24510	a Zahradníkův rok / c Karel Čapek
250	a 1. vyd.
260	a Praha : b Studio trnka, c 2009
300	a 108 s. : b il. (převážně barev.) ; c 23 cm
4901	a Edice naší babičky
500	a Ilustrace: Joseph-Pierre Redouté
65007	a zahrádkářství 7 ph116733 2 czenas
655 7	a české prózy 7 fd133972 2 czenas
7001	a Redouté, Pierre-Joseph, d 1759-1840 7 jx20081111006 4 ill
910	a BOA001 b 2-1245.618

Obrázek 25 – Příklad záznamu ve formátu MARC 21 (záznam je převzat z [katalogu MZK](#), zde je uveden ve zkrácené formě)

2.3.2 DUBLIN CORE

Metadatový standard Dublin Core (zkráceně DC) byl vytvořen v roce 1995 ve městě Dublin (v americkém státě Ohio) a spravuje ho organizace [Dublin Core Metadata Initiative](#). Původně sloužil pro popis webových objektů, následně se díky své jednoduchosti začal používat i v prostředí knihoven pro popis digitálních dokumentů (protože standard MARC byl příliš složitý pro jejich popis). Slovní vyjádření prvků a jednoduchá struktura zajišťují velkou srozumitelnost standardu.

Dublin Core je tvořen těmito 15 metadatovými prvky: Title (Název), Creator (Tvůrce), Subject (Předmět a klíčová slova), Description (Popis), Publisher (Vydavatel), Contributor (Příspěvatel), Date (Datum), Type (Typ zdroje), Format (Formát), Identifier (Identifikátor zdroje), Source (Zdroj), Language (Jazyk), Relation (Vztah), Coverage (Pokrytí), Rights (Správa autorských práv).

Podrobný výklad jednotlivých prvků je na stránkách [Dublin Core Czech](#). Každý prvek je nepovinný, opakovatelný a nemá přesně určené pořadí. Dublin Core se dá používat dvěma způsoby – s kvalifikátory nebo bez. U nekvalifikovaného Dublin Core se

používá pouze uvedených 15 prvků. Kvalifikovaný Dublin Core navíc používá tzv. kvalifikátory – specifikační prvky, které rozšiřují a zpřesňují základní elementy (např. `Creator.personalName` a `Creator.corporateName`).

```
<dc>
  <dc:title>Zahradníkův rok</dc:title>
  <dc:title>Bylinkový kalendář báby kořenářky 2010</dc:title>
  <dc:creator>Čapek, Karel</dc:creator>
  <dc:subject>zahradkářství</dc:subject>
  <dc:subject>821.162.3-3</dc:subject>
  <dc:description>Ilustrace: Joseph-Pierre Redouté</dc:description>
  <dc:publisher>Studio trnka</dc:publisher>
  <dc:contributor>Redouté, Pierre-Joseph</dc:contributor>
  <dc:date>2009</dc:date>
  <dc:type>model:monograph</dc:type>
  <dc:format>108 s. : il. (převážně barev.) ; 23 cm</dc:format>
  <dc:format>print</dc:format>
  <dc:identifier>urnnbn:urn:nbn:cz:mzk-005i1f</dc:identifier>
  <dc:identifier>URN:ISBN:978-80-87209-32-5</dc:identifier>
  <dc:identifier>uuid:b4bcc350-44f7-11e7-a34b-005056827e51</dc:identifier>
  <dc:identifier>barCode:2610443352</dc:identifier>
  <dc:identifier>ccnb:cnb001997242</dc:identifier>
  <dc:language>cze</dc:language>
  <dc:rights>policy:private</dc:rights>
</dc>
```

Obrázek 26 – Příklad záznamu v Dublin Core. Záznam můžete porovnat se záznamem v MARC 21.

2.3.3 MODS

[MODS](#) (Metadata Object Description Schema) je metadatové schéma, které bylo vytvořeno za účelem popisu digitálních dokumentů v digitálních knihovnách. Schéma vzniklo v Kongresové knihovně v roce 2002 a nejnovější verze 3.7 pochází z roku 2018. Umožňuje podrobnější popis než Dublin Core a zároveň je jednodušší než MARC 21. Schéma MODS je založené na jazyce XML, takže pomocí šablon je možné převádět data z formátů [MARC 21](#) či [Dublin Core](#) do formátu MODS.

MODS obsahuje 20 základních popisných prvků, které se dále dělí na dílčí prvky (podprvky). K prvkům i podprvkům lze přidat atributy. Prvky i podprvky jsou nepovinné a opakovatelné a podprvky mají pevně dané pořadí. Atributy jsou zase neopakovatelné a nemají pevně dané pořadí. Stejně jako u Dublin Core jsou názvy prvků a podprvků v MODS vyjádřeny slovně, takže na první pohled je MODS srozumitelnější než MARC 21.

Mezi základní prvky patří `titleInfo`, `name`, `typeOfResource`, `genre`, `originInfo`, `language`, `physicalDescription`, `abstract`, `tableOfContents`, `targetAudience`, `note`, `subject`, `classification`, `relatedItem`, `identifier`, `location`, `accessCondition`, `part`, `extension` a `recordInfo`.

Všechny prvky, podprvky a atributy jsou podrobně popsány na [webových stránkách standardu MODS](#).

```
<mods>
  <mods:titleInfo>
    <mods:title>Zahradníkův rok</mods:title>
  </mods:titleInfo>
  <mods:titleInfo type="alternative">
    <mods:title>Bylinkový kalendář báby kořenářky 2010</mods:title>
  </mods:titleInfo>
  <mods:name type="personal" usage="primary">
    <mods:namePart>Čapek, Karel</mods:namePart>
    <mods:namePart type="date">1890-1938</mods:namePart>
  </mods:name>
  <mods:name type="personal">
    <mods:namePart>Redouté, Pierre-Joseph</mods:namePart>
    <mods:namePart type="date">1759-1840</mods:namePart>
  </mods:name>
  <mods:typeOfResource>text</mods:typeOfResource>
  <mods:originInfo>
    <mods:place>
      <mods:placeTerm type="text">Praha</mods:placeTerm>
    </mods:place>
    <mods:publisher>Studio trnka</mods:publisher>
    <mods:edition>1. vyd.</mods:edition>
    <mods:dateIssued>2009</mods:dateIssued>
  </mods:originInfo>
  <mods:language>
    <mods:languageTerm authority="iso639-2b"
type="code">cze</mods:languageTerm>
  </mods:language>
  <mods:physicalDescription>
    <mods:form authority="marcform">print</mods:form>
    <mods:extent>108 s. : il. (převážně barev.) ; 23 cm</mods:extent>
  </mods:physicalDescription>
  <mods:note>Ilustrace: Joseph-Pierre Redouté</mods:note>
  <mods:subject authority="czenas">
    <mods:topic>zahradkářství</mods:topic>
  </mods:subject>
  <mods:classification authority="udc" edition="Konspekt">821.162.3-
3</mods:classification>
  <mods:location>
    <mods:physicalLocation authority="siglaADR">BOA001</mods:physicalLocation>
    <mods:shelfLocator>2-1245.618</mods:shelfLocator>
  </mods:location>
  <mods:identifier type="urnnbn">urn:nbn:cz:mzk-005i1f</mods:identifier>
  <mods:identifier type="barCode">2610443352</mods:identifier>
  <mods:identifier type="uuid">b4bcc350-44f7-11e7-a34b-
005056827e51</mods:identifier>
  <mods:identifier type="ccnb">cnb001997242</mods:identifier>
  <mods:identifier type="isbn">978-80-87209-32-5</mods:identifier>
  <mods:identifier type="isbn">978-80-87209-33-2</mods:identifier>
  <mods:recordInfo>
    <mods:descriptionStandard>aacr</mods:descriptionStandard>
  </mods:recordInfo>
</mods>
```

Obrázek 27 – Příklad záznamu v MODS. Záznam můžete porovnat se záznamy MARC 21 a Dublin Core.

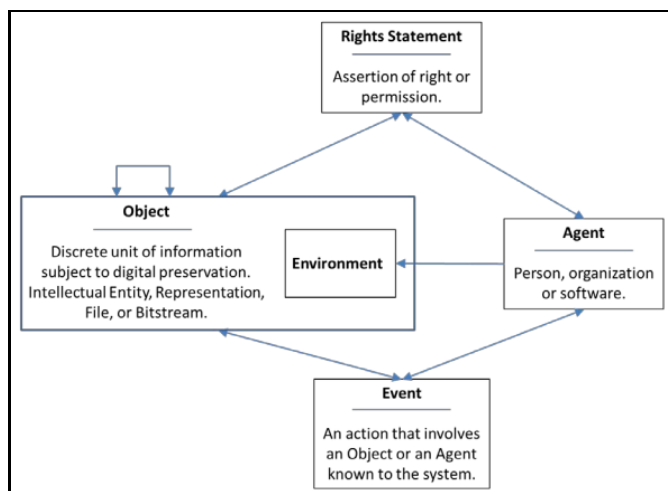
2.4 Administrativní metadata

Administrativní metadata **obsahují informace potřebné ke správě digitálního objektu nebo informace vztahující se ke vzniku objektu**. Administrativní metadata zahrnují technická metadata, archivační metadata a právní metadata.¹⁶

- **Technická metadata** popisují technické vlastnosti digitálního objektu. Mohou to být informace o souborech (formát souboru, velikost souboru, barevnost a rozlišení u skenů), o použitých softwarech či o skenerech.
- **Archivační metadata** (anglicky Preservation metadata) podporují dlouhodobou archivaci a ochranu digitálních objektů. Zahrnují informace o vzniku digitálního objektu a informace o změnách, ke kterým došlo během digitalizace nebo uchovávání.
- **Právní metadata** (anglicky Rights metadata) poskytují informace týkající se duševního vlastnictví digitálního objektu (licence, držitelé práva, datum možného zveřejnění atd.).

2.4.1 PREMIS

PREMIS (PREservation Metadata Implementation Strategies) je standard spravovaný Kongresovou knihovnou určený **pro dlouhodobé uchování digitálních dokumentů**. Zahrnuje archivační, technická i právní metadata, která zajišťují dlouhodobou použitelnost a přístupnost digitálních objektů. [Standard](#) obsahuje schéma, slovník a dokumentaci. Datový model PREMIS definuje 4 entity, které jsou důležité z hlediska digitálního uchovávání. Jsou to objekty, události, agenti a práva. Více informací k PREMIS naleznete [zde](#), příklad PREMIS záznamu se nachází [zde](#).



Obrázek 28 – Datový model PREMIS¹⁷

¹⁶ RILEY, Jenn. *Understanding Metadata: What is metadata and what is it for?* [online]. 2017, s. 6 [cit. 2021-9-29]. ISBN 978-1-937522-72-8. Dostupné z:

https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf

¹⁷ PREMIS: *Data Dictionary for Preservation Metadata* [online]. Version 3.0. PREMIS Editorial Committee, 2015 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>

2.4.2 MIX

[MIX](#) (Metadata for Images in XML Schema) je XML schéma spravované Kongresovou knihovnou, které slouží **pro zápis technických metadat obrazových souborů**. Pomocí toho schématu můžeme u obrazových souborů popsat velikost, rozměry, formát, verzi formátu, kompresi, rozlišení, barevný profil či informace o použitém skeneru. Příklad záznamu v MIX se nachází [zde](#).

2.4.3 TEXTMD

[TextMD](#) (Technical Metadata for Text) je schéma, které podrobně popisuje **technická metadata pro textové digitální dokumenty**. TextMD umožňuje popsat kódování znaků, znakové sady, fonty, velikost písma, zalomení řádků, jazyk, technické požadavky pro tisk a prohlížení, řazení stránek a [mnoho dalších údajů](#). TextMD často slouží jako rozšiřující schéma v METS (viz kapitola METS) v sekci administrativních metadat či v PREMIS. Opět se jedná o schéma zapsané v XML a spravované Kongresovou knihovnou. Na příklad TextMD záznamu se můžete podívat [zde](#).

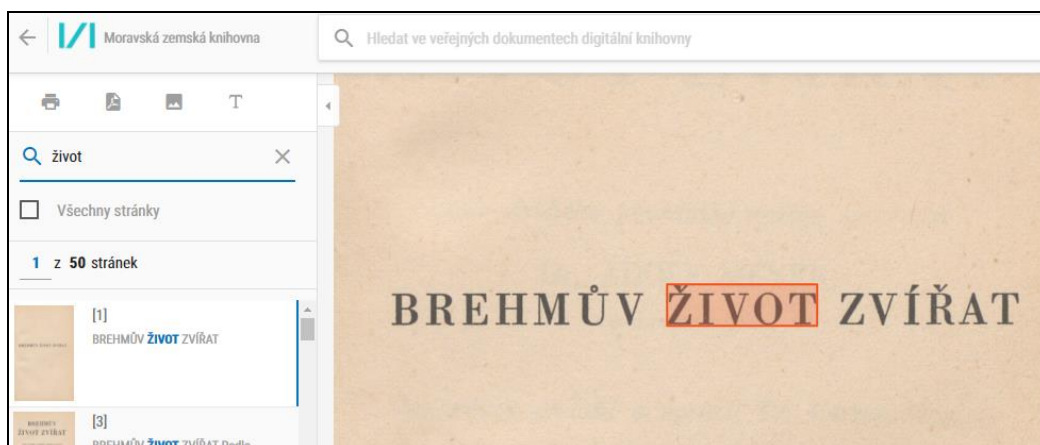
2.4.4 ALTO

[ALTO](#) (Analyzed Layout and Text Object) je metadatový formát, který **zaznamenává rozložení textu a obsah textu v digitálním dokumentu**. Informace o textu jsou získávány pomocí technologie OCR¹⁸. ALTO určuje přesnou polohu a hodnotu jednotlivých částí na stránce (ať už článků, odstavců anebo jednotlivých slov). Díky tomu je možné digitální dokument fulltextově prohledávat (uživatel se zobrazí přesná poloha slova na skenu). ALTO je spravované Kongresovou knihovnou a je zapsané v XML.

ALTO se skládá z 3 částí:

- **Description** (Popis) – Informace o samotném ALTO souboru + informace o tom, jak byl soubor vytvořen.
- **Styles** (Styly) – Informace o použitých fontech (typ, velikost, barva atd.) a vzhledu (zarovnání odstavců).
- **Layout** (Rozvržení) – Informace o obsahu dokumentu. Jsou zde například uvedeny rozměry strany či rozměry okrajů, poloha odstavců a řádků a pak samozřejmě jednotlivá slova včetně přesné polohy. Jednotlivé sekce mohou mít své identifikátory, na které může být odkazováno v METS dokumentu (viz kapitola METS). Bližší informace k rozvržení naleznete [zde](#).

¹⁸ OCR (Optical Character Recognition = optické rozpoznání znaků) je technologie, pomocí které lze převádět obrazová data (skeny) do počítačem čitelné textové podoby. Úspěšnost OCR je závislá především na kvalitě předlohy. U nových dokumentů je úspěšnost OCR vysoká, u starších dokumentů či rukopisů bývá úspěšnost nižší. V současné době vznikají OCR nástroje, které využívají nejnovější poznatky strojového učení a zpracování jazyka a které by měly být schopné rozpoznat i rukopisy. Příkladem je projekt [PERO](#) (společný projekt MZK a VUT).



Obrázek 30 – ALTO záznam umožňuje fulltextové vyhledávání v digitálním dokumentu. Vyhledané slovo se zvýrazní, aby byla patrná jeho poloha.

2.5 Strukturální metadata

Pomocí strukturálních metadat můžeme **vyjádřit vnitřní organizaci a strukturu digitálního objektu**. Digitální objekt totiž často obsahuje více částí, jako například skeny ve více verzích (uživatelská a archivní kopie), soubory s OCR a popisná či jiná metadata. Díky strukturálním metadatům můžeme tyto části smysluplně provázat. Dále můžeme popsat dělení digitálního objektu do menších celků — u periodik můžeme popsat ročníky, čísla a články, u monografií můžeme popsat kapitoly. Ve strukturálních metadatech je také zapsané, jak jdou jednotlivé strany za sebou a o jaké typy stran se jedná (titulní list, obsah, rejstřík atd.). Všechny tyto údaje zajišťují smysluplné zobrazování dokumentu v digitální knihovně a snadnou navigaci v dokumentu. Zároveň jsou tyto údaje užitečné pro výměnu dokumentů mezi institucemi či pro jejich dlouhodobé uchovávání.

2.5.1 METS

METS (Metadata Encoding & Transmission Standard) je standard, který slouží k popisu struktury digitálního objektu. K obrazovým datům připojuje popisná, administrativní a strukturální metadata. Je to jakýsi kontejner, do kterého mohou být vloženy různé metadatové formáty (MODS, DC, ALTO, MIX). METS se zapisuje do XML a je spravován Kongresovou knihovnou. Struktura METS dokumentu se skládá ze sedmi hlavních sekcí:

1. **Hlavička METS** <metsHdr> – Obsahuje popis samotného METS dokumentu (datum vytvoření, tvůrce, editor atd.).
2. **Sekce popisných metadat** <dmdSec> – Obsahuje popisná metadata ve formátu MODS, DC, MARCXML apod.
3. **Sekce administrativních metadat** <amdSec> – Obsahuje administrativní metadata (jak technická, právní, tak i archivační metadata) ve formátu PREMIS, MIX, TextMD apod.
4. **Sekce souborů** <fileSec> – Obsahuje seznam souborů, které jsou součástí digitálního objektu. Soubory mohou být rozdělené do skupin.

5. **Strukturální mapa** <structMap> – Nejdůležitější sekce METS (jediná povinná sekce). Vyjadřuje hierarchickou strukturu digitálního objektu a propojuje skeny s metadaty. Obsahuje fyzickou mapu (kde se nachází soubory) a logickou mapu (jaký k sobě mají vztah).
6. **Strukturální odkazy** <structLink> – Vyjadřují vazby mezi fyzickou a logickou strukturální mapou.
7. **Sekce pravidel chování** <behaviorSec> – Tato sekce definuje chování mezi jednotlivými objekty v METS.

Na příklady METS záznamů se můžete podívat [zde](#).

2.6 Standardy Národní digitální knihovny (NDK)

Pro projekt Národní digitální knihovna byly vytvořeny **metadatové standardy (tzv. NDK standardy), které byly následně využity i v krajských digitalizacích a které jsou povinně předepsány v rámci programu [VISK 7](#)**. Momentálně existují standardy pro monografie, periodika, některé zvukové dokumenty (gramofonové desky, fonografické válečky) a pro e-publikace. Dále vznikl metodický pokyn pro digitalizaci starých tisků a rukopisů, který ale vychází ze standardu pro monografie. Do budoucna by měly vzniknout standardy i pro další druhy dokumentů (například pro kroniky či pro zvuková CD). Standardy vytváří a spravuje Oddělení standardů v Národní knihovně ČR.

Ve standardech NDK jsou využity metadatové formáty, které jsme si popsali v předchozích kapitolách. Výsledkem digitalizace dle standardu NDK by měl být balíček dat a metadat (tzv. NDK balíček), který obsahuje tyto části:

- popisná metadata (MODS, Dublin Core),
- administrativní metadata (PREMIS),
- technická metadata (MIX, AES57 u zvukových souborů),
- strukturální metadata (METS),
- archivní kopie skenů (bezztrátový JPEG2000 u obrazových souborů, WAV u zvukových souborů, PDF/A a EPUB pro e-publikace),
- uživatelské kopie skenů (ztrátový JPEG2000 u obrazových souborů, MP3 u zvukových souborů, PDF/A a EPUB pro e-publikace),
- soubory s OCR (ALTO, TXT),
- kontrolní metadatové soubory (s kontrolními součty a údaji o vzniku dat).

Aktuální verze standardů se nachází na webové stránce [Národní digitální knihovny](#). Na této stránce jsou k dispozici i vzorové NDK balíčky, které si můžete stáhnout a projít.

2.6.1 FORMÁTOVÝ VÝBOR NDK

Formátový výbor NDK je **poradní orgán Národní knihovny ČR pro oblast metadat a souborových formátů digitálních dat**. Formátový výbor rozhoduje o potřebě aktualizace stávajících standardů či o vzniku nových standardů. Vedle Formátového

výboru ještě existují pracovní skupiny, které jsou výboru podřízené. V současné době existují tři: pro digitalizaci textových dokumentů, pro audiovizuální dokumenty a pro e-publikace. Do pracovních skupin se může přihlásit kterýkoliv pracovník digitalizace v ČR, jehož instituce používá NDK standardy. Více informací k Formátovému výboru NDK se nachází [zde](#).

SHRNUTÍ STUDIJNÍHO TEXTU

Digitalizace umožňuje knihovnám a dalším paměťovým institucím snadné zpřístupňování fondů uživatelům a také slouží k dlouhodobému uchovávání kulturního dědictví. Při skenování, zpřístupňování i při archivaci dokumentů se používají různé druhy identifikátorů a digitálním objektům jsou přiřazována metadata.

Identifikátory slouží k rozlišení jednotlivých entit v určitém systému. V knihovnách se používají různé typy identifikátorů; pro digitalizaci jsou podstatné ty, které popisují knihovní dokumenty, konkrétně provedení díla (ISBN, ISSN, ISMN, čČNB) a jednotky (čárový kód). Při digitalizaci se identifikátory používají také k provázání fyzické knihy s digitální kopií. Pomocí dalších identifikátorů pak rozlišujeme digitální objekty (UUID, URN, URN:NBN, Handle, DOI, ARK) a případně k nim poskytujeme i přístup (URL, PURL).

Metadata jsou informace, pomocí kterých popisujeme digitální objekty a které nám usnadňují jejich vyhledávání, používání a spravování. Abychom byli při vytváření metadat jednotní, musíme si vytvořit metadatové schéma či použít již existující standardizované schéma. Mezinárodní metadatové standardy jsou potřebné i pro snadnou výměnu dat mezi institucemi. Metadatové standardy bývají zapsané ve značkovacím jazyku XML (XML se dělí do částí, kterým se říká elementy, a každý element má počáteční a ukončovací tag). Metadata dělíme na popisná, administrativní a strukturální. Popisná metadata identifikují daný objekt a usnadňují jeho vyhledávání, mezi nejpoužívanější standardy patří MODS a Dublin Core. Administrativní metadata zaznamenávají informace potřebné ke správě digitálního objektu a mohou obsahovat technická, archivační a právní metadata. Do administrativních metadat řadíme standardy PREMIS, MIX, TextMD a ALTO. Strukturální metadata vyjadřují vnitřní organizaci a strukturu digitálního objektu, příkladem je standard METS.

LITERATURA

- ALTO: Technical Metadata For Layout and Text Objects. *Library of Congress* [online]. 2016 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/alto/>
- ALTO: Technical Metadata For Layout and Text Objects: ALTO <Styles> & <Layout> Usage. *Library of Congress* [online]. 2016 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/alto/techcenter/layout.html>
- ARK. *Confluence: CDL Curation* [online]. 22. 3. 2018 [cit. 2021-10-5]. Dostupné z: <https://confluence.ucop.edu/display/Curation/ARK>
- ARK Alliance: Home of the Archival Resource Key (ARK) [online]. [cit. 2021-10-5]. Dostupné z: <https://arks.org/>
- ARK NAANs and systems. *ARK Alliance: Home of the Archival Resource Key (ARK)* [online]. [cit. 2021-10-5]. Dostupné z: <https://arks.org/about/ark-naans-and-systems/>
- BALÍKOVÁ, Marie. Identifikátor (indexace). In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2021-10-01]. Dostupné z: https://aleph.nkp.cz/F/?func=direct&doc_number=000001548&local_base=KTD
- BELLINI, Emanuele, Chiara CIRINNA a Maurizio LUNGHI. Trvalé identifikátory pro kulturní dědictví. *Digital Preservation Europe* [online]. 2008 [cit. 2021-10-01]. Dostupné z: https://issuu.com/dpe-staff/docs/cz_trvale_identifikatory
- CUBR, Ladislav a Zdeněk VAŠEK. Identifikátory digitálních dokumentů se zaměřením na systém URN:NBN v ČR. *Čtenář: Měsíčník pro knihovny* [online]. 2013, 65(6) [cit. 2021-10-03]. Dostupné z: <https://svkkl.cz/en/ctenar/clanek/1254>
- CZIDLO: CZech IDentification and LOcalization tool [online]. Národní knihovna České republiky [cit. 2021-10-3]. Dostupné z: <https://resolver.nkp.cz/web/#tab0>
- Čárové kódy pro české knihovny. *Národní technická knihovna* [online]. [cit. 2021-7-28]. Dostupné z: https://www.techlib.cz/cs/2920-carove-kody-pro-ceske-knihovny#tab_howto
- České národní středisko ISSN. *Národní technická knihovna* [online]. [cit. 2021-7-28]. Dostupné z: <https://www.techlib.cz/cs/2844-ceske-narodni-stredisko-issn>
- Číslo České národní bibliografie (čČNB) a jak ho získat. *Registr digitalizace* [online]. [cit. 2021-7-28]. Dostupné z: <http://www.registrdigitalizace.cz/rdcz/info/data/ccnb>
- Číslo ČNB v SK ČR. *Souborný katalog České republiky* [online]. [cit. 2021-7-28]. Dostupné z: <https://www.caslin.cz/caslin/spoluprace/sluzby-souborneho-katalogu-cr/cislo-cnb-v-sk-cr>
- Detailed Description of MODS Elements. *The Library of Congress* [online]. 2009 [cit. 2021-9-29]. Dostupné z: <https://www.loc.gov/standards/mods/v3/mods-userguide-elements.html>

Digital object identifier. *Wikipedie* [online]. 2021 [cit. 2021-10-3]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=Digital_object_identifier&oldid=19452419

DOI® [online]. 2021 [cit. 2021-10-3]. Dostupné z: <https://www.doi.org/>

Dublin Core Czech [online]. [cit. 2021-9-29]. Dostupné z: http://webserver.ics.muni.cz/dublin_core/index.html

Dublin Core Metadata Element Set Mapping to MODS Version 3. *The Library of Congress* [online]. 2012 [cit. 2021-9-29]. Dostupné z: <https://www.loc.gov/standards/mods/dcsimple-mods.html>

Dublin Core Metadata Initiative [online]. [cit. 2021-9-29]. Dostupné z: <https://www.dublincore.org/>

Extensible Markup Language. *Wikipedie* [online]. [cit. 2021-8-23]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=Extensible_Markup_Language&oldid=20379729

EZID. *California Digital Library: University of California* [online]. 8. 11. 2019 [cit. 2021-10-5]. Dostupné z: <https://cdlib.org/services/uc3/ezid/>

Fees. *CrossRef* [online]. 2019 [cit. 2021-10-3]. Dostupné z: <https://www.crossref.org/fees/>

Formátový výbor NDK. *Národní digitální knihovna* [online]. Národní knihovna ČR, 10.03.2021 [cit. 2021-9-30]. Dostupné z: <https://standardy.ndk.cz/ndk/formatovy-vybor-ndk>

Funkční požadavky na bibliografické záznamy. Přeložila Ludmila Celbová. Praha: Národní knihovna České republiky, 2001. 115 s. ISBN 80-7050-400-5. Dostupné z: <http://archive.ifla.org/VII/s13/frbr/frbr-cs.pdf>

International Standard Book Number. *Wikipedie* [online]. [cit. 2021-7-28]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=International_Standard_Book_Number&oldid=19929658

International Standard Serial Number. *Wikipedie* [online]. [cit. 2021-7-28]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=International_Standard_Serial_Number&oldid=19978440

JANOVSKÝ, Dušan. URL. *Jak psát web* [online]. [cit. 2021-8-20]. Dostupné z: <https://www.jakpsatweb.cz/html/url.html>

JÍLKOVÁ, Marta. *Persistentní identifikátory a jejich využití v digitálních knihovnách* [online]. Praha, 2013 [cit. 2021-9-4]. Dostupné z: <https://dspace.cuni.cz/bitstream/handle/20.500.11956/51726/120138608.pdf?sequence=1&isAllowed=y>. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta. Vedoucí práce Barbora Drobíková.

Kaskádové styly. *Wikipedie* [online]. [cit. 2021-9-29]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=Kask%C3%A1dov%C3%A9_styly&oldid=19856484

KOSEK, Jiří. *Základy jazyka XML* [online]. [cit. 2021-8-23]. Dostupné z: <https://www.kosek.cz/clanky/swn-xml/syntaxe.html>

MARC 21. *Národní knihovna České republiky* [online]. 2019 [cit. 2021-9-29]. Dostupné z: <https://www.nkp.cz/o-knihovne/odborne-cinnosti/zpracovani-fondu/katalogizacni-politika/katalogizacni-politika/marc21-soubor>

MARC 21 Format for Bibliographic data. *Library of Congress* [online]. 2021 [cit. 2021-9-29]. Dostupné z: <https://www.loc.gov/marc/bibliographic/>

MARC 21 to MODS 3.7 Mapping. *The Library of Congress* [online]. 2020 [cit. 2021-9-29]. Dostupné z: <https://www.loc.gov/standards/mods/v3/mods-mapping-3-7.html>

MARCXML: MARC 21 XML Schema. *Library of Congress* [online]. 2020 [cit. 2021-9-29]. Dostupné z: <http://www.loc.gov/standards/marcxml/>

Metadata Object Description Schema: Official Web Site. *The Library of Congress* [online]. 2020 [cit. 2021-9-29]. Dostupné z: <https://www.loc.gov/standards/mods/>

METS: Metadata Encoding & Transmission Standard. *Library of Congress* [online]. 2020 [cit. 2021-9-30]. Dostupné z: <http://www.loc.gov/standards/mets/>

METS: Metadata Encoding & Transmission Standard: METS Example Documents. *Library of Congress* [online]. 2016 [cit. 2021-9-30]. Dostupné z: <http://www.loc.gov/standards/mets/mets-examples.html>

Mezinárodní registrační systémy. *Národní knihovna České republiky* [online]. 20.02.2021 [cit. 2021-7-28]. Dostupné z: <https://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn>

MIX: Metadata For Images in XML Schema: Technical Metadata For Digital Still Images Standard. *Library of Congress* [online]. 2020 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/mix/>

NAAN Request Form. *ARK Alliance* [online]. [cit. 2021-10-5]. Dostupné z: <https://docs.google.com/forms/d/e/1FAIpQLSfd1CX6idwLB47g8OGKUG654auV8IU8yI7DAs61cXGOoFDn0g/viewform?c=0&w=1>

NISO. *Understanding Metadata* [online]. 2004, s. 1 [cit. 2021-9-27]. ISBN 1-880124-62-9. Dostupné z: https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf

NOID: Nice Opaque IDentifier (minter and name resolver). *N2t: RESOLVING Names to Things* [online]. [cit. 2021-10-5]. Dostupné z: <https://n2t.net/e/noid.html>

PREMIS: *Data Dictionary for Preservation Metadata* [online]. Version 3.0. PREMIS Editorial Committee, 2015 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>

PREMIS: Preservation Metadata Maintenance Activity. *Library of Congress* [online]. 2021 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/premis/index.html>

PREMIS: Preservation Metadata Maintenance Activity: PREMIS Data Dictionary for Preservation Metadata, Version 3.0. *Library of Congress* [online]. 2020 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/premis/v3/index.html>

PREMIS XML Usage Examples. *Library of Congress* [online]. 2017 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/premis/examples.html>

Project PERO [online]. [cit. 2021-10-26]. Dostupné z: <https://pero.fit.vutbr.cz/>

Přidělování DOI. *Univerzita Tomáše Bati ve Zlíně: Knihovna* [online]. [cit. 2021-10-3]. Dostupné z: <https://knihovna.utb.cz/veda-a-vyzkum/elektronicke-publikovani/pridelovani-doi/>

RDA. *Národní knihovna České republiky* [online]. 2021 [cit. 2021-9-29]. Dostupné z: <https://www.nkp.cz/o-knihovne/odborne-cinnosti/zpracovani-fondu/katalogizacni-politika/rda>

Resolve a DOI Name. *DOI®* [online]. [cit. 2021-10-3]. Dostupné z: <https://dx.doi.org/>

Resolve a Handle and View the Values. *Handle.Net* [online]. [cit. 2021-10-3]. Dostupné z: <http://hdl.handle.net/>

RILEY, Jenn. *Understanding Metadata: What is metadata and what is it for?* [online]. 2017, s. 6 [cit. 2021-9-29]. ISBN 978-1-937522-72-8. Dostupné z: https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf

RYCHTÁŘOVÁ, Pavla, KRAVECOVÁ, Lenka a Markéta KRUTSKÁ: *Identifikátory*. E-learningový kurz Digitalizace v knihovnách [online]. Brno: Moravská zemská knihovna v Brně, 2021 [cit. 2021-8-22]. Dostupné z: <https://kurzy.knihovna.cz/>

RYCHTÁŘOVÁ, Pavla, KRAVECOVÁ, Lenka a Markéta KRUTSKÁ: *Metadata I*. E-learningový kurz Digitalizace v knihovnách [online]. Brno: Moravská zemská knihovna v Brně, 2021 [cit. 2021-8-22]. Dostupné z: <https://kurzy.knihovna.cz/>

RYCHTÁŘOVÁ, Pavla, KRAVECOVÁ, Lenka a Markéta KRUTSKÁ: *Metadata II*. E-learningový kurz Digitalizace v knihovnách [online]. Brno: Moravská zemská knihovna v Brně, 2021 [cit. 2021-8-22]. Dostupné z: <https://kurzy.knihovna.cz/>

Standardy pro metadata. *Národní digitální knihovna* [online]. Národní knihovna ČR, 13.01.2021 [cit. 2021-9-30]. Dostupné z: <https://standardy.ndk.cz/ndk/standardy-digitalizace/metadata>

TextMD: Technical Metadata for Text. *Library of Congress* [online]. 2021 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/textMD/>

TextMD: Technical Metadata for Text: The textMD v.3.0 alpha element set. *Library of Congress* [online]. 2021 [cit. 2021-9-30]. Dostupné z: <https://www.loc.gov/standards/textMD/elementSet/index.html>

Uniform Resource Identifier. *Wikipedie* [online]. [cit. 2021-9-4]. Dostupné z: https://cs.wikipedia.org/wiki/Uniform_Resource_Identifier

Uniform Resource Identifier (URI) Schemes. *Iana* [online]. [cit. 2021-9-4]. Dostupné z: <https://www.iana.org/assignments/uri-schemes/uri-schemes.xhtml>

Uniform Resource Name. *Wikipedie* [online]. [cit. 2021-9-4]. Dostupné z: https://cs.wikipedia.org/wiki/Uniform_Resource_Name

Uniform Resource Names (URN) Namespaces. *Iana* [online]. [cit. 2021-9-5]. Dostupné z: <https://www.iana.org/assignments/urn-namespaces/urn-namespaces.xml>

Univerzální unikátní identifikátor. *Wikipedie* [online]. [cit. 2021-7-28]. Dostupné z: https://cs.wikipedia.org/w/index.php?title=Univerz%C3%A1ln%C3%AD_unik%C3%A1tn%C3%AD_identifik%C3%A1tor&oldid=19787907

UTF-8. *Wikipedie* [online]. [cit. 2021-9-29]. Dostupné z: <https://cs.wikipedia.org/w/index.php?title=UTF-8&oldid=20417874>

What Is an ISMN? *International ISMN Agency* [online]. [cit. 2021-7-28]. Dostupné z: <https://www.ismn-international.org/whatis>

XML Tutorial. *W3 Schools* [online]. [cit. 2021-8-23]. Dostupné z: <https://www.w3schools.com/xml/default.asp>

XML Schema. *Wikipedie* [online]. [cit. 2021-8-24]. Dostupné z: https://cs.wikipedia.org/wiki/XML_Schema

ZEDNÍČEK, Jan. SQL XML – Základy XML – Struktura, tagy, elementy a atributy (Díl 2.). *Biportal* [online]. [cit. 2021-8-23]. Dostupné z: <https://biportal.cz/sql-zaklady-xml-tagy-elementy/>

Název:	Orientace ve standardech digitalizace
Verze:	1.0
Autor:	Mgr. Markéta Krutská
Počet stran:	49
Rok vydání:	2021
Jazyková korektura:	Markéta Roupcová