

Univerzita Karlova v Praze
Filozofická fakulta
Ústav informačních studií a knihovnictví

Informační studia a knihovnictví – Informační věda

Mgr. Linda Jansová

**Citování vědecké literatury prostřednictvím
formátu XML**

**Citing Scientific Literature Using an XML
Format**

Disertační práce

vedoucí práce – Ing. Miloslav Nič, Ph.D.

2010

Prohlášení:

Prohlašuji, že jsem disertační práci vykonala samostatně s využitím uvedených pramenů a literatury.

V Praze, 30. června 2010

.....

podpis

Identifikační záznam:

JANSOVÁ, Linda. *Citování vědecké literatury prostřednictvím formátu XML [Citing Scientific Literature Using an XML Format]*. Praha, 2010. 245 s., 61 s. příl. Disertační práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Školitel Ing. Miloslav Nič, Ph.D.

Abstrakt:

Disertace je zaměřena na problematiku využití jazyka XML pro účely citování. Nejprve je přiblížen současný stav problematiky citování, který je významně ovlivněn rozvojem elektronického publikování. Následně jsou představeny vybrané projekty zaměřené na strojově srozumitelná citační data a způsoby zahrnutí citačních dat do XML formátů. Podrobně jsou charakterizovány citační formáty užívané v praxi, a to jak formáty využívající jazyk XML, tak další formáty. Stručně jsou představeny vybrané citační manažery jako nástroje využívané při automatizované práci s citačními daty. Pozornost je věnována také typologii citovaných dokumentů. Zahrnuta je i podrobná analýza citační praxe uplatňované v tzv. barevných knihách IUPAC. Na základě dílčích zjištění byly formulovány doporučené postupy pro zahrnutí citačních dat do XML formátů. Doporučené postupy mají podobu sedmi zásad, které lze stručně shrnout takto: využití kódování UTF-8, nízký počet prvků i atributů, maximální strukturování dat, využití řízených slovníků, zachycení hierarchie dat a propojení dat pomocí odkazů, využití existující typologie dokumentů nebo vytvoření nové typologie, využití principu rekurze při práci s daty. Na základě doporučených postupů byl navržen experimentální formát. Možnost jeho implementace v praxi je ukázána na příkladu citačních dat z barevných knih IUPAC.

Abstract:

The dissertation focuses on the use of XML for citation purposes. First, a current situation in citation practice, deeply influenced by the development in electronic publishing, is described. Then selected projects aimed at making citation data machine understandable and ways to include citation data into XML formats are presented. Citation formats (both XML and other formats) are dealt with in detail.

This part is followed by a brief introduction of citation managers as tools used to automate work with citation data. Typology of cited documents is also covered. Furthermore, a detailed analysis of citation practice in IUPAC color books is included. Based on partial findings recommended practices for the inclusion of citation data into XML formats were developed. The recommended practices are a set of seven principles which can be summed up as follows: use of UTF-8 encoding, low number of elements and attributes, maximum data structuring, use of controlled vocabularies, data hierarchy and use of links to make connections, use of existing document typology or creation of a new one, use of recursion while working with the data. An experimental format based on the recommended practices has been designed. Citation data from IUPAC color books are used to present the way how the format can be implemented in practice.

Klíčová slova:

citační formáty, citační manažery, citování, elektronické publikování, IUPAC, typologie dokumentů, XML formáty

Keywords:

citation formats, citation managers, citing, electronic publishing, IUPAC, document typology, XML formats

Předmluva

Předkládaná disertační práce je zaměřena na oblast citování vědecké literatury prostřednictvím formátu XML¹. Vzhledem ke skutečnosti, že v současné době dochází k výraznému rozvoji v oblasti elektronického publikování, lze toto téma pokládat za vysoce aktuální.

Automatické propojování dokumentů prostřednictvím citačních dat je žádoucí mj. s ohledem na dynamicky se rozvíjející oblast digitálních knihoven a informačních zdrojů obdobného charakteru. Za účelem automatického propojování dat je však nezbytným předpokladem mít k dispozici data, na jejichž základě lze toto propojování provádět. Vědci a odborníci ve svých pracích využívají řadu různých citačních stylů, s nimiž není vždy snadné dále pracovat bez zásahu lidské ruky, tj. s využitím softwaru. Odlišnosti ve způsobech citování lze nalézt jak na úrovni syntaktické (zejména interpunkce v záznamech citovaných dokumentů), tak na úrovni sémantické (konkrétní údaje uváděné v záznamech). Zejména se sémantickými odlišnostmi se pak lze setkat i v případě citačních formátů určených primárně pro zachycení citačních dat a jejich import do citačních manažerů, popř. export z těchto softwarů.

Při práci na disertačním projektu jsem využila zejména své zkušenosti z oblasti elektronického publikování a redakční práce – v letech 2005-2010 jsem působila na pozici výkonné redaktorky elektronického časopisu Ikaros. Zároveň jsem navázala na poznatky získané z projektů, na nichž jsem se podílela jako informační specialistka (v některých případech jako koordinátorka), přičemž součástí mých povinností bylo mj. provádět rešerše a zpracovávat přehledy relevantních zdrojů. Dalším zdrojem přínosných poznatků se staly výběrové semináře věnované citování a částečně též výběrové semináře zaměřené na knihovní software Evergreen (semináře byly realizovány v letech 2007-2010 na Ústavu informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze).

¹ XML je rozšiřitelný značkovací jazyk (*Extensible Markup Language*), který tvoří podmnožinu univerzálního značkovacího metajazyka SGML (*Standard Generalized Markup Language*). Jeho cílem je umožnit využití metajazyka SGML ve webovém prostředí [1], [2].

Cíle disertační práce byly vytvořit doporučené postupy pro zahrnutí citačních dat do XML formátů se zaměřením na oblast STM² a implementovat je v publikačních systémech IUPAC³.

Aby bylo možné výše stanovené cíle splnit, byl zpracován přehled způsobů zahrnutí citačních dat do vybraných existujících formátů z oblasti STM (doplněný o charakteristiku vybraných projektů souvisejících s oblastí strojově strojitelných citačních dat), dále přehled vybraných citačních formátů (využívajících i nevyužívajících jazyk XML) a přehled vybraných citačních manažerů (dostupných zdarma a/nebo s otevřeným zdrojovým kódem) se zaměřením na formáty pro import a export dat. Doplnkově byla také připravena analýza vybraných citačních stylů se zaměřením na nejlépe hodnocené vědecké časopisy a typologie citovaných dokumentů. S ohledem na předpokládané využití XML formátu v publikačních systémech IUPAC byl také proveden průzkum citační praxe v tzv. barevných knihách IUPAC.

Na základě výstupů z výše zmíněných kroků byly vypracovány doporučené postupy pro zahrnutí citačních dat do XML formátů a byl zpracován návrh experimentálního formátu. Ten byl následně prakticky uplatněn při zpracování citačních dat ze tří z tzv. barevných knih IUPAC. Zpracovaná data byla také využita k vizualizaci citačních vztahů.

Výsledky práce jsou prezentovány v sedmi hlavních kapitolách, které následují po úvodu, v němž jsou podrobněji komentovány jednotlivé kroky zpracování disertačního projektu. Kapitola 2 se zabývá současným stavem problematiky citování z vybraných hledisek – prezentuje mj. zmiňované výsledky analýzy citačních stylů v nejlépe hodnocených vědeckých časopisech. Téma disertačního projektu tak zasazuje do širšího kontextu. Kapitola 3 je věnována strojově srozumitelným citačním datům, jsou v ní stručně představeny vybrané projekty a způsoby zahrnutí citačních dat do vybraných XML formátů. Kapitola 4 analyzuje a hodnotí devět vybraných citačních formátů, kapitola 5 se zabývá citačními manažery. Kapitola 6 poskytuje přehled vybraných typologií dokumentů, které se vyskytují v analyzovaných citačních formátech a dále ve vybraných citačních stylech a ontologiích; doplněny jsou také typologie dokumentů z oblasti

² STM je zkratka výrazu *Science, Technology, Medicine* (věda, technika, medicína).

³ IUPAC je zkratka názvu *International Union of Pure and Applied Chemistry* (Mezinárodní unie pro čistou a užitou chemii).

knihovnictví. V kapitole 7 jsou obsaženy výsledky analýzy citační praxe v barevných knihách IUPAC. Kapitola 8 vychází z dílčích závěrů předchozích kapitol. Právě na jejich základě jsou formulovány doporučené postupy pro zahrnutí citačních dat do XML formátů a je navržen experimentální formát. V kapitole 9 jsou uvedeny celkové závěry.

Hlavní text práce doplňují seznam obrázků, zdrojů obrázků, tabulek a zkratek a slovník vybraných termínů, dále pak seznam použitých pramenů a literatury a přílohy. Součástí textu je také poznámkový aparát, který je uváděn formou poznámek pod čarou. Poznámkový aparát obsahuje rovněž výklad jednotlivých zkratek v prvním místě jejich výskytu.

Formální úprava disertace vychází z pokynů Filozofické fakulty Univerzity Karlovy v Praze a z normy ČSN ISO 7144.

Použité prameny a literatura jsou citovány v souladu s normami ČSN ISO 690 a ČSN ISO 690-2, pro citace v textu je použita číselná metoda (v případech, kdy je citována konkrétní část daného dokumentu, např. konkrétní stránka, následují tyto údaje po číslu odkazujícím na položku seznamu použitých pramenů a literatury).

Výsledky uvedené v této práci byly již částečně publikovány – část kapitol 2 a 7 byla publikována v časopise ITlib [3], kapitoly 4 a 5 zčásti využívají poznatků zveřejněných v příspěvcích připravených do sborníku ze semináře CASLIN⁴ [4], [5].

Za poskytnutí cenných rad a konzultací v průběhu zpracovávání práce bych ráda poděkovala svému školiteli Miloslavu Ničovi z Vysoké školy chemicko-technologické v Praze. Bez jeho trpělivosti a vstřícnosti by se mi disertační práci nepodařilo úspěšně dokončit. Poděkování dále patří Bedřichu Košatovi z téže instituce, kolegyním a kolegům z Ústavu informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze, pracovníkům Národní knihovny České republiky (zejména Evě Mackové z Knihovny knihovnické literatury a Marii Balíkové z Oddělení věcných autorit a věcného zpracování) a Knihovny společenských věd T. G. Masaryka v Jinonicích (zejména Aleně Matuszkové). V neposlední řadě bych ráda vyjádřila dík svým rodičům a svému manželovi za podporu a vytváření zázemí po celou dobu studia.

⁴ CASLIN je zkratka výrazu *Czech and Slovak Library Information Network* (Česká a slovenská knihovní informační síť).

Obsah

1. Úvod.....	1
2. Současný stav problematiky citování.....	4
2.1 Úvod.....	4
2.2 Citování ve vztahu k teorii informační vědy a knihovnicko-informační praxi	4
2.3 Citování jako součást informační gramotnosti.....	6
2.4 Standardizace v oblasti citování.....	6
2.5 Citační styly v nejlépe hodnocených vědeckých časopisech.....	8
2.6 Automatizované přebírání citačních dat.....	17
2.7 Proměna vztahů mezi aktéry informační sféry společnosti.....	18
2.8 Elektronické publikování jako změna paradigmatu publikování.....	19
2.9 Dílčí závěry.....	22
3. Strojům srozumitelná citační data.....	24
3.1 Úvod.....	24
3.2 Vybrané projekty.....	24
3.2.1 CSL.....	24
3.2.2 BIBO.....	26
3.2.3 CiTO.....	28
3.2.4 OOoBib.....	30
3.3 Citační data ve vybraných XML formátech.....	31
3.3.1 DocBook.....	32
3.3.2 OpenDocument Format for Office Applications.....	36
3.3.3 EPUB.....	39
3.3.4 TEI.....	40
3.3.5 CML.....	42
3.3.6 ThermoML.....	43
3.3.7 Interní XML formát Zlaté knihy.....	45
3.4 Dílčí závěry.....	47
4. Citační formáty.....	48
4.1 Úvod.....	48
4.2 Tvorba datových sad.....	51
4.3 Bílé znaky.....	57
4.4 Formáty nevyužívající jazyk XML.....	59
4.4.1 RIS.....	59
4.4.2 ISI.....	69
4.4.3 RefWorks	74
4.4.4 EndNote.....	79
4.4.5 BibTeX	89
4.4.6 OpenURL COinS.....	100
4.5 Formáty využívající jazyk XML.....	104
4.5.1 MARCXML.....	104
4.5.2 MODS.....	106
4.5.3 NLM.....	111
4.6 Dílčí závěry.....	116

5. Citační manažery.....	118
5.1 Úvod.....	118
5.2 Zotero.....	119
5.3 Connotea.....	122
5.4 Citace 2.0.....	126
5.5 CiteULike.....	128
5.6 Dílčí závěry.....	130
6. Typologie citovaných dokumentů.....	131
6.1 Úvod.....	131
6.2 Termín „dokument“.....	131
6.3 Přehled vybraných zdrojů věnovaných typologiím dokumentů.....	133
6.4 Typologie dokumentů v citačních formátech.....	136
6.4.1 RIS.....	136
6.4.2 ISI.....	137
6.4.3 RefWorks.....	139
6.4.4 EndNote.....	141
6.4.5 BibTeX.....	144
6.4.6 OpenURL COinS.....	146
6.4.7 MARCXML.....	147
6.4.8 MODS.....	148
6.4.9 NLM.....	149
6.5 Typologie dokumentů v citačních stylech.....	151
6.5.1 APA.....	151
6.5.2 ACS.....	152
6.5.3 ISO 690 a 690-2.....	154
6.6 Typologie dokumentů v citačních ontologiích.....	156
6.6.1 CiTO.....	156
6.6.2 BIBO.....	159
6.7 Typologie dokumentů v oblasti knihovnictví.....	161
6.7.1 AACR2.....	161
6.7.2 RDA.....	162
6.7.3 MDT.....	164
6.8 Dílčí závěry.....	165
7. Citační praxe v tzv. barevných knihách IUPAC.....	167
7.1 Úvod.....	167
7.2 IUPAC a publikační činnost.....	167
7.3 Barevné knihy.....	170
7.4 Zlatá kniha – Chemická terminologie.....	170
7.4.1 Záznamy použité literatury a citace.....	171
7.4.2 Bibliografické údaje o knize.....	172
7.5 Zelená kniha – Množství, jednotky a symboly ve fyzikální chemii.....	173
7.5.1 Záznamy použité literatury a citace.....	173
7.5.2 Bibliografické údaje o knize.....	175
7.6 Červená kniha – Nomenklatura anorganické chemie.....	176
7.6.1 Záznamy použité literatury a citace.....	176
7.6.2 Bibliografické údaje o knize.....	178
7.7 Modrá kniha – Nomenklatura organických sloučenin.....	179

7.7.1 Záznamy použité literatury a citace.....	179
7.7.2 Bibliografické údaje o knize.....	180
7.8 <i>Fialová kniha – Makromolekulární nomenklatura</i>	181
7.8.1 Záznamy použité literatury a citace.....	182
7.8.2 Bibliografické údaje o knize.....	185
7.9 <i>Oranžová kniha – Analytická nomenklatura</i>	186
7.9.1 Záznamy použité literatury a citace.....	186
7.9.2 Bibliografické údaje o knize.....	189
7.10 <i>Stříbrná kniha – Nomenklatura a symboly v klinické chemii</i>	191
7.11 <i>Bílá kniha – Biochemická nomenklatura a související dokumenty</i>	191
7.11.1 Záznamy použité literatury a citace.....	191
7.11.2 Bibliografické údaje o knize.....	195
7.12 <i>Dílčí závěry</i>	197
8. Zahrnutí citačních dat do XML formátů.....	199
8.1 <i>Úvod</i>	199
8.2 <i>Doporučené postupy</i>	200
8.3 <i>Návrh experimentálního formátu</i>	203
8.3.1 <i>Návrhová kritéria</i>	203
8.3.2 <i>Metodika přípravy</i>	204
8.3.3 <i>Praktické využití formátu na příkladu citačních dat z barevných knih</i>	207
8.3.4 <i>Vizualizace citačních vztahů</i>	215
8.4 <i>Dílčí závěry</i>	217
9. Závěr.....	219
Seznam použitých pramenů a literatury.....	223
Přílohy.....	1
<i>Příloha č. 1: Web of Science</i>	1
<i>Příloha č. 2: Scopus</i>	5
<i>Příloha č. 3: Nature</i>	8
<i>Příloha č. 4: Science</i>	9
<i>Příloha č. 5: SpringerLink</i>	13
<i>Příloha č. 6: ScienceDirect</i>	14
<i>Příloha č. 7: Wiley InterScience</i>	15
<i>Příloha č. 8: PubMed</i>	18
<i>Příloha č. 9: EBSCOhost</i>	21
<i>Příloha č. 10: ProQuest Central</i>	25
<i>Příloha č. 11: WorldCat</i>	28
<i>Příloha č. 12: bibliotek.dk</i>	29
<i>Příloha č. 13: CERN Document Server</i>	30
<i>Příloha č. 14: E-LIS</i>	36
<i>Příloha č. 15: GOLX-3-47</i>	48
<i>Příloha č. 16: GOLX-3-58</i>	50
<i>Příloha č. 17: GOLX-3-93</i>	52
<i>Příloha č. 18: GOLX-3-103</i>	54
<i>Příloha č. 19: RED3-13-13A</i>	56
<i>Příloha č. 20: RED3-14-7D</i>	58
<i>Příloha č. 21: RED3-45-1C</i>	60

Seznam obrázků

Obr. 1: Ukázka části obohaceného seznamu použitých zdrojů v článku budoucnosti z časopisu Cell.....	21
Obr. 2: Část stylu APA (6th ed.) v CSL.....	26
Obr. 3: Část vizualizace ontologie BIBO.....	27
Obr. 4: Část vizualizace ontologie CiTO.....	29
Obr. 5: Zachycení citačních vztahů v preprintu CiTO, the Citation Typing Ontology D. Shottona.....	30
Obr. 6: Záznam časopiseckého článku v ontologii BIBO.....	31
Obr. 7: Panel zásuvného modulu Zotera v textovém editoru OpenOffice.org Writer	31
Obr. 8: Část příkladu schématu Relax NG formátu DocBook týkající se atributu „enumeration“.....	35
Obr. 9: Část průvodce věnovaná prvku „author“.....	36
Obr. 10: Pole pro zachycení položky použité literatury v editoru OpenOffice.org Writer.....	38
Obr. 11: Část příkladu užití prvků „listBibl“, „biblStruct“, „bibl“ a dalších z dokumentace verze P5 formátu TEI.....	41
Obr. 12: Příklad example1.xml ze specifikace CML.....	43
Obr. 13: Příklad údajů o knize v interním XML formátu Zlaté knihy.....	46
Obr. 14: Část záznamu časopiseckého článku z digitální knihovny ScienceDirect ve formátu ASCII.....	49
Obr. 15: Záznam časopiseckého článku ze systému PubMed ve formátu PubMed Citation.....	49
Obr. 16: Zobrazení záznamu knihy ze sady RIS-11.....	62
Obr. 17: Záznam časopiseckého článku ze sady ISI-1.....	71
Obr. 18: Část prvního záznamu ze sady v XML verzi formátu RefWorks.....	78
Obr. 19: Část prvního záznamu ze sady v neaktuálním XML verzi formátu RefWorks.....	79
Obr. 20: Část DTD pro online verzi softwaru EndNote se zvýrazněním rozdílů oproti DTD pro desktopovou verzi softwaru.....	85
Obr. 21: Část záznamu ze sady v exportním XML formátu softwaru EndNote.....	88
Obr. 22: Část záznamu e-printu ve formátu Zotero RDF.....	120
Obr. 23: Část záznamu e-printu ve formátu MODS.....	120

Obr. 24: Část záznamu e-printu ve formátu Unqualified Dublin Core RDF.....	121
Obr. 25: Úvodní část souboru s daty uloženými citačním manažerem Zotero.....	121
Obr. 26: Část dotazu v jazyce SQL pro vytvoření tabulky „itemTypes“.....	122
Obr. 27: Část záznamu exportovaného ve formátu MODS (XML).....	123
Obr. 28: Část záznamu exportovaného ve formátu Word 2007 (Bibliography).....	124
Obr. 29: Část záznamu exportovaného ve formátu RDF.....	124
Obr. 30: Část záznamu exportovaného ve formátu RSS.....	125
Obr. 31: Část dotazu v jazyce SQL pro vytvoření tabulky „citation“.....	125
Obr. 32: Import záznamů do citačního manažeru Citace.com.....	127
Obr. 33: Import záznamů do citačního manažeru CiteULike.....	129
Obr. 34: Typologie dokumentů ve formátu RIS.....	136
Obr. 35: Typologie dokumentů v systému Web of Science.....	138
Obr. 36: Typologie dokumentů ve formátu RefWorks.....	140
Obr. 37: Typologie dokumentů ve formátu EndNote.....	142
Obr. 38: Typologie dokumentů ve formátu BibTeX.....	144
Obr. 39: Typologie knih a jejich částí ve formátu info:ofi/fmt:kev:mtx:book.....	146
Obr. 40: Typologie časopiseckých a příbuzných zdrojů ve formátu info:ofi/fmt:kev:mtx:journal.....	146
Obr. 41: Typologie dokumentů ve formátu MARCXML.....	147
Obr. 42: Typologie dokumentů ve formátu MODS ve verzi 3.3.....	148
Obr. 43: Část typů dokumentů ve formátu NLM.....	150
Obr. 44: Typologie dokumentů v publikačním manuálu APA (5th ed.).....	151
Obr. 45: Typologie dokumentů v publikačním manuálu APA (6th ed.).....	152
Obr. 46: Typologie dokumentů v publikačním manuálu ACS (2nd ed.).....	153
Obr. 47: Typologie dokumentů v publikačním manuálu ACS (3rd ed.).....	154
Obr. 48: Typologie dokumentů v normě ISO 690.....	155
Obr. 49: Typologie dokumentů v normě ISO 690-2.....	155
Obr. 50: Typologie dokumentů v ontologii CiTO – úroveň Work (dílo).....	157
Obr. 51: Typologie dokumentů v ontologii CiTO – úroveň Expression (vyjádření)	158
Obr. 52: Typologie dokumentů v ontologii CiTO – úroveň Manifestation (provedení)	158
.....	
Obr. 53: Typologie dokumentů v ontologii BIBO.....	160

Obr. 54: Jedna z typologií dokumentů v Anglo-amerických katalogizačních pravidlech.....	161
Obr. 55: Typy obsahu v pravidlech RDA.....	163
Obr. 56: Typologie znaků formy v MDT.....	164
Obr. 57: Rozcestník k publikacím organizace IUPAC.....	168
Obr. 58: Příklad použitých a dalších zdrojů.....	172
Obr. 59: Vzor záznamu konkrétního termínu ve Zlaté knize.....	172
Obr. 60: Část zdrojového kódu stránky z webové prezentace Zlaté knihy.....	173
Obr. 61: Příklad abecedně-číselných citací v textu.....	174
Obr. 62: Údaje na titulní straně Zelené knihy.....	175
Obr. 63: Bibliografické údaje o Zelené knize ve vlastnostech dokumentu ve formátu PDF.....	176
Obr. 64: Příklad záznamu s poznámkou o zkráceném označení dokumentu a plánovaném novém vydání.....	177
Obr. 65: Příklad citací použité literatury z kapitoly „IR-1 General Aims, Functions and Methods of Chemical Nomenclature“ z Červené knihy.....	178
Obr. 66: Údaje na titulní straně Červené knihy.....	178
Obr. 67: Bibliografické údaje o Červené knize ve vlastnostech dokumentu ve formátu PDF.....	179
Obr. 68: Odkazovaný záznam.....	179
Obr. 69: Příklad citace a záznamu v textu.....	180
Obr. 70: Část zdrojového kódu stránky z webové prezentace Modré knihy.....	181
Obr. 71: Bibliografické údaje vztahující se k jednotlivým částem Fialové knihy....	185
Obr. 72: Část zdrojového kódu domácí stránky webové prezentace Fialové knihy.	186
Obr. 73: Ukázka ze seznamu zdrojů s uvedením vztahu k jednotlivým částem kapitoly.....	188
Obr. 74: Příklad záznamů v textu Oranžové knihy.....	189
Obr. 75: Bibliografické údaje o tištěné verzi Oranžové knihy na obálce.....	189
Obr. 76: Zdrojový kód hlavního rámce domácí stránky online verze Oranžové knihy	190
Obr. 77: Metadata v souboru ve formátu PDF s kapitolou 7.4 z Oranžové knihy....	190
Obr. 78: Bibliografické údaje týkající se tištěného vydání Bílé knihy.....	195
Obr. 79: Část zdrojového kódu domácí stránky webové prezentace Bílé knihy.....	195

Obr. 80: Část zdrojového kódu jedné ze stránek webové prezentace Bílé knihy.....	196
Obr. 81: Bibliografické údaje o Bílé knize ve vlastnostech dokumentu ve formátu PDF.....	196
Obr. 82: Ukázka souboru ve formátu DOT.....	215
Obr. 83: Vizualizace části vztahů mezi dokumenty.....	216
Obr. 84: Vizualizace všech identifikovaných vztahů mezi dokumenty.....	217

Seznam zdrojů obrázků

Poznámky:

V případech, kdy není konkrétní zdroj obrázku či podkladových dat citován v textu, je uveden pouze formou adresy a data získání. Je-li naproti tomu zdroj citován i jinde v textu, je uvedena citace odkazující na položku v seznamu použitých pramenů a literatury.

Zdroje obrázků v přílohách jsou zároveň zdroji, z nichž byly získávány datové sady v kapitole 4.

V případech, kdy se jedná o licencované zdroje, jedná se převážně o adresy funkční pouze v prostředí Univerzity Karlovy v Praze.

Zdroje obrázků v hlavní textové části

1. <http://www.cell.com/abstract/S0092-8674%2809%2901439-1> [získáno 2010-06-11], zobrazení v internetovém prohlížeči Mozilla Firefox
2. <http://www.zotero.org/styles/apa> [získáno 2010-06-02], zobrazení v XML editoru oXygen
3. [6], zobrazení v internetovém prohlížeči Mozilla Firefox
4. <http://bibliontology.com/content/article> [získáno 2010-06-02], zobrazení v internetovém prohlížeči Mozilla Firefox
5. [7], zobrazení v internetovém prohlížeči Mozilla Firefox
6. [8], zobrazení v internetovém prohlížeči Mozilla Firefox, resp. v aplikaci Adobe Acrobat Reader
7. OpenOffice.org Writer [získáno 2010-06-02]
8. <http://docs.oasis-open.org/docbook/rng/5.0/docbook.rng> [získáno 2010-05-15], zobrazení v XML editoru oXygen
9. <http://www.docbook.org/tdg5/en/html/author.html> [získáno 2010-05-15], zobrazení v internetovém prohlížeči Mozilla Firefox
10. OpenOffice.org Writer [získáno 2010-05-15]
11. <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html#COBI> [získáno 2010-05-16], zobrazení v internetovém prohlížeči Mozilla Firefox
12. <http://sourceforge.net/projects/cml/files/cml/schema2.5RC1/schema25RC1.zip/download> [získáno 2010-05-16], zobrazení v XML editoru oXygen

13. [9], zobrazení v XML editoru oXygen
14. <http://www.sciencedirect.com.ezproxy.is.cuni.cz/science> [získáno 2010-03-16], zobrazení v internetovém prohlížeči Mozilla Firefox
15. <http://www.ncbi.nlm.nih.gov/pubmed/20110493?dopt=Abstract> [získáno 2010-02-04], zobrazení v internetovém prohlížeči Mozilla Firefox
16. Sada RIS-11, zobrazení v textovém editoru PSPad
17. Sada ISI-1, zobrazení v textovém editoru PSPad
18. <http://www.refworks.com.ezproxy.is.cuni.cz/> [získáno 2010-04-07], zobrazení v XML editoru oXygen
19. <http://www.refworks.com.ezproxy.is.cuni.cz/> [získáno 2010-04-07], zobrazení v XML editoru oXygen
20. [10] [získáno 2010-04-07], zobrazení v softwaru KDiff3 pro porovnávání souborů
21. [11], zobrazení v XML editoru oXygen
22. Zotero [získáno 2010-05-16], zobrazení v XML editoru oXygen
23. Zotero [získáno 2010-05-16], zobrazení v XML editoru oXygen
24. Zotero [získáno 2010-05-16], zobrazení v XML editoru oXygen
25. Zotero, soubor zotero.sqlite [získáno 2010-05-16], zobrazení v textovém editoru PSPad
26. [12], zobrazení v internetovém prohlížeči Mozilla Firefox
27. Connotea [získáno 2010-06-03], zobrazení v XML editoru oXygen
28. Connotea [získáno 2010-06-03], zobrazení v XML editoru oXygen
29. Connotea [získáno 2010-06-03], zobrazení v XML editoru oXygen
30. Connotea [získáno 2010-06-03], zobrazení v XML editoru oXygen
31. [13, soubor schema.sql], zobrazení v programátorském editoru Bluefish
32. <http://citace.com/hledat.php?ukol=1&nazev=&prijmeni=Cejpek&isbn=&odeslat=Vyhledat> [získáno 2010-06-03], zobrazení v internetovém prohlížeči Mozilla Firefox
33. [14], zobrazení v internetovém prohlížeči Mozilla Firefox
34. [15], grafické zpracování v softwaru FreeMind
35. [16], grafické zpracování v softwaru FreeMind
36. [17], grafické zpracování v softwaru FreeMind

37. [18, s. 711-739 (Appendix C Reference Types and Fields)], grafické zpracování v softwaru FreeMind
38. [19, s. 14-15], grafické zpracování v softwaru FreeMind
39. [20], grafické zpracování v softwaru FreeMind
40. [21], grafické zpracování v softwaru FreeMind
41. [22], grafické zpracování v softwaru FreeMind
42. [23], grafické zpracování v softwaru FreeMind
43. [24], grafické zpracování v softwaru FreeMind
44. [25], grafické zpracování v softwaru FreeMind
45. [26], grafické zpracování v softwaru FreeMind
46. [27], grafické zpracování v softwaru FreeMind
47. [28], grafické zpracování v softwaru FreeMind
48. [29], grafické zpracování v softwaru FreeMind
49. [30], grafické zpracování v softwaru FreeMind
50. [31], grafické zpracování v softwaru FreeMind
51. [31], grafické zpracování v softwaru FreeMind
52. [31], grafické zpracování v softwaru FreeMind
53. [6], grafické zpracování v softwaru FreeMind
54. [32], grafické zpracování v softwaru FreeMind
55. [33], grafické zpracování v softwaru FreeMind
56. [34], grafické zpracování v softwaru FreeMind
57. <http://www.iupac.org/Publications> [získáno 2009-08-26], zobrazení v internetovém prohlížeči Mozilla Firefox
58. <http://goldbook.iupac.org/P04524.html> [získáno 2009-08-31], zobrazení v internetovém prohlížeči Mozilla Firefox
59. <http://goldbook.iupac.org/P04524.html> [získáno 2009-08-31], zobrazení v internetovém prohlížeči Mozilla Firefox
60. <http://goldbook.iupac.org/P04524.html> [získáno 2009-09-02], zobrazení v internetovém prohlížeči Mozilla Firefox
61. [35, s. viii], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
62. [35], zobrazení v aplikaci Adobe Acrobat Reader
63. [35], zobrazení metadat v aplikaci Adobe Acrobat Reader

64. [36 s. 45], zobrazení v aplikaci Adobe Acrobat Reader
65. [36, s. 2], zobrazení v aplikaci Adobe Acrobat Reader
66. [36], zobrazení v aplikaci Adobe Acrobat
67. [36], zobrazení metadat v aplikaci Adobe Acrobat Reader
68. http://www.acdlabs.com/iupac/nomenclature/93/r93_615.htm [získáno 2009-08-30], zobrazení v internetovém prohlížeči Mozilla Firefox
69. http://www.acdlabs.com/iupac/nomenclature/93/r93_36.htm [získáno 2009-08-30], zobrazení v internetovém prohlížeči Mozilla Firefox
70. [37], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
71. [38] , zobrazení v internetovém prohlížeči Mozilla Firefox
72. [38], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
73. http://media.iupac.org/publications/analytical_compendium/Cha01refs.pdf [získáno 2009-08-30]. zobrazení v aplikaci Adobe Acrobat Reader
74. http://media.iupac.org/publications/analytical_compendium/Cha11sec1.pdf [získáno 2009-09-03], zobrazení v aplikaci Adobe Acrobat Reader
75. [39], zobrazení v internetovém prohlížeči Mozilla Firefox
76. http://media.iupac.org/publications/analytical_compendium/CoverPage.html [získáno 2009-09-02], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
77. http://media.iupac.org/publications/analytical_compendium/Cha07sec4.pdf [získáno 2009-09-02], zobrazení metadat v aplikaci Adobe Acrobat Reader
78. [40], zobrazení v internetovém prohlížeči Mozilla Firefox
79. [40], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
80. <http://www.chem.qmul.ac.uk/iupac/misc/pnuc1.html> [získáno 2009-09-03], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
81. <http://www.iupac.org/publications/pac/1983/pdf/5508x1273.pdf> [získáno 2009-09-03], zobrazení metadat v aplikaci Adobe Acrobat Reader
82. Soubor DOT [získáno 2010-05-31], zobrazení v textovém editoru PSPad
83. Graphviz [získáno 2010-05-31], zobrazení v grafickém editoru Inkscape
84. Graphviz [získáno 2010-05-31], zobrazení v grafickém editoru Inkscape

Zdroje obrázků v přílohách

Příloha č. 1:

1. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
2. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
3. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v internetovém prohlížeči Mozilla Firefox
4. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
5. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
6. <http://apps.isiknowledge.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad

Příloha č. 2:

1. <http://www.scopus.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení
v textovém editoru PSPad
2. <http://www.scopus.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení
v textovém editoru PSPad
3. <http://www.scopus.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení
v textovém editoru PSPad
4. <http://www.scopus.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení
v textovém editoru PSPad

Příloha č. 3:

1. <http://www.nature.com/nature/journal/v463/n7281/ris/nature08729.ris>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
2. [http://www.connotea.org/add?
uri=http://www.nature.com/nature/journal/v463/n7281/full/nature08729.html](http://www.connotea.org/add?uri=http://www.nature.com/nature/journal/v463/n7281/full/nature08729.html)
[získáno 2010-02-04], zobrazení v internetovém prohlížeči Mozilla Firefox

Příloha č. 4:

1. <http://www.sciencemag.org/cgi/citmgr?type=endnote&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad
2. <http://www.sciencemag.org/cgi/citmgr?type=refman&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad
3. <http://www.sciencemag.org/cgi/citmgr?type=procite&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad
4. <http://www.sciencemag.org/cgi/citmgr?type=bibtex&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad
5. <http://www.sciencemag.org/cgi/citmgr?type=refworks&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad
6. <http://www.sciencemag.org/cgi/citmgr?type=medlars&gca=sci;323/5910/54>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
7. <http://www.sciencemag.org/cgi/citmgr?type=zotero&gca=sci;323/5910/54>
[získáno 2010-04-04], zobrazení v textovém editoru PSPad

Příloha č. 5:

1. <http://www.springerlink.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
2. <http://www.springerlink.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad

Příloha č. 6:

1. <http://www.sciencedirect.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
2. <http://www.sciencedirect.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad
3. <http://www.sciencedirect.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04],
zobrazení v textovém editoru PSPad

Příloha č. 7:

1. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad

2. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad
3. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad
4. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad
5. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad
6. <http://www3.interscience.wiley.com.ezproxy.is.cuni.cz/> [získáno 2010-04-04], zobrazení v textovém editoru PSPad

Příloha č. 8:

1. <http://www.ncbi.nlm.nih.gov/pubmed/20204640> [získáno 2010-04-05], zobrazení v textovém editoru PSPad
2. <http://www.ncbi.nlm.nih.gov/pubmed/20204640> [získáno 2010-04-05], zobrazení v textovém editoru PSPad
3. <http://www.ncbi.nlm.nih.gov/pubmed/20204640> [získáno 2010-04-05], zobrazení v textovém editoru PSPad
4. <http://www.ncbi.nlm.nih.gov/pubmed/20204640> [získáno 2010-04-05], zobrazení v XML editoru oXygen
5. <http://www.ncbi.nlm.nih.gov/pubmed/20204640> [získáno 2010-04-05], zobrazení v textovém editoru PSPad

Příloha č. 9:

1. <http://web.ebscohost.com/> [získáno 2010-04-05], zobrazení v internetovém prohlížeči Mozilla Firefox
2. <http://web.ebscohost.com/> [získáno 2010-04-05], zobrazení v textovém editoru PSPad
3. <http://web.ebscohost.com/> [získáno 2010-04-05], zobrazení v XML editoru oXygen
4. <http://web.ebscohost.com/> [získáno 2010-04-05], zobrazení v internetovém prohlížeči Mozilla Firefox

5. <http://web.ebscohost.com/> [získáno 2010-04-05], zobrazení v textovém editoru PSPad

Příloha č. 10:

1. <http://proquest.umi.com.ezproxy.is.cuni.cz/> [získáno 2010-04-05], zobrazení v internetovém prohlížeči Mozilla Firefox
2. <http://proquest.umi.com.ezproxy.is.cuni.cz/> [získáno 2010-04-05], zobrazení v textovém editoru PSPad
3. <http://proquest.umi.com.ezproxy.is.cuni.cz/> [získáno 2010-04-05], zobrazení v textovém editoru PSPad

Příloha č. 11:

1. http://www.worldcat.org/title/digital-libraries/oclc/45731551&referer=brief_results [získáno 2010-04-05], zobrazení v internetovém prohlížeči Mozilla Firefox
2. http://www.worldcat.org/title/digital-libraries/oclc/45731551&referer=brief_results [získáno 2010-04-05], zobrazení v textovém editoru PSPad

Příloha č. 12:

1. <http://bibliotek.dk/> [získáno 2010-04-05], zobrazení v textovém editoru PSPad

Příloha č. 13:

1. <http://cdsweb.cern.ch/record/1214893/export/hx> [získáno 2010-02-04], zobrazení v textovém editoru PSPad
2. <http://cdsweb.cern.ch/record/1214893/export/hm> získáno 2010-02-04], zobrazení v textovém editoru PSPad
3. <http://cdsweb.cern.ch/record/1214893/export/xm> [získáno 2010-02-04], zobrazení v editoru oXygen
4. <http://cdsweb.cern.ch/record/1214893/export/xd> [získáno 2010-02-04], zobrazení v editoru oXygen

5. <http://cdsweb.cern.ch/record/1214893/export/xo> [získáno 2010-02-04],
zobrazení v editoru oXygen
6. <http://cdsweb.cern.ch/record/1214893/export/xn> [získáno 2010-02-04],
zobrazení v editoru oXygen

Příloha č. 14:

1. <http://eprints.rclis.org/cgi/export/17305/HTML/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení v internetovém prohlížeči Mozilla Firefox
2. <http://eprints.rclis.org/cgi/export/17305/HTML/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení zdrojového kódu v internetovém prohlížeči
Mozilla Firefox
3. <http://eprints.rclis.org/cgi/export/17305/Text/eprints3-eprint-17305.txt>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
4. <http://eprints.rclis.org/cgi/export/17305/Full/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení v internetovém prohlížeči Mozilla Firefox
5. <http://eprints.rclis.org/cgi/export/17305/Full/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení zdrojového kódu v internetovém prohlížeči
Mozilla Firefox
6. <http://eprints.rclis.org/cgi/export/17305/ContextObject/eprints3-eprint-17305.xml> [získáno 2010-02-04], zobrazení v XML editoru oXygen
7. <http://eprints.rclis.org/cgi/export/17305/EndNote/eprints3-eprint-17305.enw>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
8. <http://eprints.rclis.org/cgi/export/17305/BibTeX/eprints3-eprint-17305.bib>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
9. <http://eprints.rclis.org/cgi/export/17305/COinS/eprints3-eprint-17305.txt>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
10. <http://eprints.rclis.org/cgi/export/17305/MODS/eprints3-eprint-17305.xml>
[získáno 2010-02-04], zobrazení v XML editoru oXygen
11. <http://eprints.rclis.org/cgi/export/17305/DIDL/eprints3-eprint-17305.xml>
[získáno 2010-02-04], zobrazení v XML editoru oXygen
12. <http://eprints.rclis.org/cgi/export/17305/XML/eprints3-eprint-17305.xml>
[získáno 2010-02-04], zobrazení v XML editoru oXygen

13. <http://eprints.rclis.org/cgi/export/17305/ISO/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení v internetovém prohlížeči Mozilla Firefox
14. <http://eprints.rclis.org/cgi/export/17305/ISO/eprints3-eprint-17305.html>
[získáno 2010-02-04], zobrazení zdrojového kódu v internetovém prohlížeči Mozilla Firefox
15. <http://eprints.rclis.org/cgi/export/17305/DC/eprints3-eprint-17305.txt>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
16. <http://eprints.rclis.org/cgi/export/17305/RIS/eprints3-eprint-17305.ris>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
17. <http://eprints.rclis.org/cgi/export/17305/EAP/eprints3-eprint-17305.xml>
[získáno 2010-02-04], zobrazení v XML editoru oXygen
18. <http://eprints.rclis.org/cgi/export/17305/Simple/eprints3-eprint-17305.txt>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
19. <http://eprints.rclis.org/cgi/export/17305/Refer/eprints3-eprint-17305.refer>
[získáno 2010-02-04], zobrazení v textovém editoru PSPad
20. <http://eprints.rclis.org/cgi/export/17305/METS/eprints3-eprint-17305.xml>
[získáno 2010-02-04], zobrazení v XML editoru oXygen

Seznam tabulek

Tab. 1: Pokyny pro autory v časopisech s nejvyšším impakt faktorem.....	9
Tab. 2: Citační styly v pokynech pro autory ve vybraných časopisech.....	12
Tab. 3: Vydavatelé vybraných časopisů.....	16
Tab. 4: Přehled formátů a zdrojů dat.....	55
Tab. 5: Přehled datových sad.....	56
Tab. 6: Přehled vybraných bílých znaků.....	58
Tab. 7: Kódování souborů ve formátu RIS.....	59
Tab. 8: Oddělovací řetězce znaků ve formátu RIS.....	60
Tab. 9: Pole s názvy zdrojových dokumentů ve formátu RIS.....	64
Tab. 10: Pole s autorskými údaji ve formátu RIS.....	67
Tab. 11: Kódování souboru ve formátu ISI.....	70
Tab. 12: Oddělovací řetězce znaků ve formátu ISI.....	70
Tab. 13: Pole s názvy zdrojových dokumentů ve formátu ISI.....	72
Tab. 14: Pole s autorskými údaji ve formátu ISI.....	73
Tab. 15: Kódování souboru ve formátu RefWorks.....	74
Tab. 16: Oddělovací řetězce znaků ve formátu RefWorks.....	74
Tab. 17: Rozlišení různých typů autorů a editorů ve formátu RefWorks.....	75
Tab. 18: Pole s autorskými údaji ve formátu RefWorks.....	77
Tab. 19: Pole s názvy zdrojových dokumentů ve formátu RefWorks.....	77
Tab. 20: Kódování souboru ve formátu EndNote.....	80
Tab. 21: Oddělovací řetězce znaků ve formátu EndNote.....	81
Tab. 22: Pole s autorskými údaji ve formátu EndNote.....	83
Tab. 23: Pole s názvy zdrojových dokumentů ve formátu EndNote.....	84
Tab. 24: Prvky podřízené prvku „record“ v XML verzi formátu EndNote.....	87
Tab. 25: Prvky s atributy v XML verzi formátu EndNote.....	87
Tab. 26: Kódování souborů ve formátu BibTeX.....	90
Tab. 27: Oddělovací řetězce znaků ve formátu BibTeX.....	91
Tab. 28: Pole s názvy zdrojových dokumentů ve formátu BibTeX.....	94
Tab. 29: Pole s autorskými údaji ve formátu BibTeX.....	96
Tab. 30: Kódování souboru ve formátu OpenURL COinS.....	100
Tab. 31: Oddělovací řetězce znaků ve formátu OpenURL COinS.....	101
Tab. 32: Pole s autorskými údaji ve formátu OpenURL COinS.....	103

Tab. 33: Kódování souboru ve formátu MARCXML.....	104
Tab. 34: Kódování souboru ve formátu MODS.....	107
Tab. 35: Přehled prvků, jejich atributů a podřazených prvků ve formátu MODS....	108
Tab. 36: Kódování souboru ve formátu NLM.....	112
Tab. 37: Přehled prvků podřazených ve formátu NLM prvku „Article“, jejich atributů a podřazených prvků.....	113
Tab. 38: Prvky s názvy zdrojových dokumentů ve formátu NLM.....	114
Tab. 39: Prvky s autorskými údaji ve formátu NLM.....	115
Tab. 40: Typy dokumentů ve formátech RIS a RefWorks.....	141
Tab. 41: Typy dokumentů ve formátech RIS a EndNote.....	143
Tab. 42: Typy dokumentů ve formátech RefWorks a EndNote.....	144
Tab. 43: Typy dokumentů ve formátech BibTeX, RIS, RefWorks a EndNote.....	145
Tab. 44: Záznamy v Červené knize.....	177
Tab. 45: Doplněné odkazy na úplné texty kapitol Fialové knihy.....	181
Tab. 46: Záznamy v kapitolách Fialové knihy.....	182
Tab. 47: Citace z jednotlivých kapitol Fialové knihy.....	184
Tab. 48: Záznamy v kapitolách Oranžové knihy.....	186
Tab. 49: Záznamy v kapitolách Bílé knihy.....	191
Tab. 50: Příklad tabulky s excerpovanými citačními daty.....	205
Tab. 51: Příklad tabulky s excerpovanými citačními daty.....	207
Tab. 52: Použité prvky a atributy se svými hodnotami.....	209
Tab. 53: Příklady hodnot prvků „surname“.....	214

Seznam zkratek

AACR2	Anglo-American Cataloging Rules, 2nd Revised Edition
ACI	Autonomous Citation Indexing
ACS	American Chemical Society
AMA	American Medical Association
APA	American Psychological Association
API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
BIBO	Bibliographic Ontology
BOM	Byte Order Mark
CASLIN	Czech and Slovak Library Information Network
CASSI	Chemical Abstracts Service Source Index
CDS	CERN Document Server
CERN	Conseil Européen pour la recherche nucléaire
CiTO	Citation Typing Ontology
CML	Chemical Markup Language
CODEN	Code Number
COinS	ContextObject in SPAN
CPEP	Committee on Printed and Electronic Publications
CR	Carriage Return
CSL	Citation Style Language
DC	Dublin Core
DOI	Digital Object Identifier
DOS	Disk Operating System
DTD	Document Type Definition
E-LIS	E-prints in Library and Information Science
EBSCO	Elton B Stephens Company
EDTF	Extended Date/Time Format
EPUB	Electronic Publication
FID	Fédération Internationale de Documentation
FLUX-CiM	Flexible Unsupervised Extraction of Citation Metadata
FRBR	Functional Requirements for Bibliographic Records
GMD	General Material Designation
HT	Character Tabulation
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IBM	International Business Machines
ID	identifikátor

IDPF	International Digital Publishing Forum
IEC	International Electrotechnical Commission
IFLA	International Federation of Library Associations and Institutions
ISBD	International Standard Bibliographic Description
ISBN	International Standard Book Number
ISI	Institute for Scientific Information
ISO	International Organization for Standardization
ISSN	International Standard Serial Number
IUPAC	International Union of Pure and Applied Chemistry
JSON	JavaScript Object Notation
LCCN	Library of Congress Control Number
LF	Line Feed
MARC	Machine Readable Cataloging
MARCXML	MARC Extensible Markup Language
MDT	Mezinárodní desetinné třídění
MeSH	Medical Subject Headings
MIME	Multipurpose Internet Media Extensions
MODS	Metadata Object Description Schema
MRF	Master Reference File
NLM	National Library of Medicine
OAI-ORE	Open Archives Initiative – Object Reuse and Exchange
OASIS	Organization for the Advancement of Structured Information Standards
OCF	Open eBook Publication Structure Container Format
OEB	Open eBook Publication Structure
ONIX	ONline Information eXchange
OOoBib	OpenOffice.org Bibliographic
OPF	Open Packaging Format
OPS	Open Publication Structure
OS	Operating System
OWL	Web Ontology Language
PAC	Pure and Applied Chemistry
PDF	Portable Document Format
PHP	PHP: Hypertext Preprocessor
PMID	PubMed IDentifier / PubMed unique IDentifier
PNG	Portable Network Graphics
PRC	Palm Resource Compiler
RDA	Resource Description and Access
RDF	Resource Description Framework
Relax NG	REgular LAnguage for XML Next Generation

RIS	Research Information Systems
RSS	Really Simple Syndication
S	Space
SGML	Standard Generalized Markup Language
SQL	Structured Query Language
SRU	Search/Retrieval via URL
STM	Science, Technology, Medicine
SVG	Scalable Vector Graphics
SWAN	Semantic Web Applications in Neuromedicine
TEI	Text Encoding Initiative
ThermoML	Thermodynamics Markup Language
TRC	Thermodynamics Research Center
TXT	textfile
UDC	Universal Decimal Classification
URL	Uniform Resource Locator
UTF-8	8-bit UCS/Unicode Transformation Format
W3C	World Wide Web Consortium
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations

Slovník vybraných termínů

Ve slovníku jsou uvedeny výklady vybraných termínů užívaných v předkládané práci. Je vždy vybrán termín, který je užíván jako preferovaný, a je doplněn o překlad do angličtiny. U výkladů termínů označených hvězdičkou není uveden zdroj, neboť se jedná o výrazy v praxi užívané velmi různorodým způsobem. Výklad proto upřesňuje způsob chápání a užití těchto výrazů v této práci. U výrazů užitých v této práci ve specifickém významu je uveden výklad převzatý z externího zdroje doplněný o upřesnění způsobu užití termínu v této práci.

V případě potřeby byl podrobnější komentář k termínům zahrnut přímo do hlavního textu disertace.

Součástí práce jsou pochopitelně i termíny nezařazené do tohoto slovníku, ty však v práci nehrají klíčovou úlohu, proto jsou uvedeny pouze v hlavním textu.

bílý znak (<i>white space</i>)	(v dokumentech napsaných v jazyce XML) mezery, tabulátory a prázdné řádky, které slouží k oddělení značek za účelem zvýšení čitelnosti a většinou nejsou určeny k zahrnutí do výstupní verze dokumentu [1], [2]; v této práci jsou za bílé znaky dále považovány znaky pro nový řádek a posun na začátek řádku
citace (<i>citation</i> nebo <i>in-text citation</i>)*	stručná zmínka o použitém (citovaném) zdroji, která je součástí textu dokumentu a odkazuje na záznam s podrobnějšími údaji o tomto zdroji
citační formát (<i>citation format</i>)*	formát určený a/nebo využívaný pro zachycení citačních dat
citační manažer (<i>citation manager</i>)*	softwarový nástroj určený pro správu citačních dat
citační styl (<i>citation style</i>)*	sada pravidel pro vytváření citací a záznamů (určuje mj. typ uváděných údajů, jejich pořadí, interpunkci mezi jednotlivými údaji či způsoby typografického zvýraznění určitých údajů v záznamu)
dokument (<i>document</i>)	dokument je informační pramen, který existuje v hmotné nebo nehmotné podobě, obsahuje množinu informací a slouží k přenosu těchto informací v čase a prostoru (částečně podle [41, heslo dokument]); v této práci je jako ekvivalent používán též výraz „zdroj“
formát (<i>format</i>)	definovaná struktura pro typ informací, které jsou určeny ke zpracování, uložení na magnetická nebo optická média, zobrazení na obrazovce nebo tisku [42]
metadata (<i>metadata</i>)	strojově srozumitelné informace o webových zdrojích nebo dalších věcech [43]
ontologie (<i>ontology</i>)	formální, explicitní specifikace (sdílené) konceptualizace [44, s. 199]
rekurze (<i>recursion</i>)	proces definice či vyjádření funkce, procedury, jazykové konstrukce nebo řešení problému na základě sebe samé či samého [45]
sémantický web (<i>semantic web</i>)	web dat s významem, jemuž může počítačový program dostatečně porozumět, zjistit, co data znamenají, a podle toho je zpracovat [46]
XML (<i>XML</i>)	rozšiřitelný značkovací jazyk, který tvoří podmnožinu univerzálního značkovacího metajazyka SGML (Standard Generalized Markup Language) a jeho cílem je umožnit využití metajazyka SGML ve webovém prostředí [1], [2]
záznam (<i>reference</i>)*	soubor údajů o citovaném zdroji potřebný pro nalezení úplného textu zdroje a typicky uvedený nikoliv v textu, ale v rámci seznamu použitých zdrojů; termín „záznam“ je v této práci použit i ve významu anglického výrazu <i>record</i>

1. Úvod

Při plnění cílů disertačního projektu, jehož výsledky jsou prezentovány v předkládané práci, tj. vytvoření doporučených postupů pro zahrnutí citačních dat do XML formátů v oblasti STM a jejich implementace v publikačních systémech IUPAC, jsem podnikla řadu kroků, z nichž ty nejvýznamnější jsou uvedeny v následujících odstavcích úvodu. Podrobnosti o zvolených metodikách jsou potom zařazeny do jednotlivých kapitol vztahujících se přímo ke zpracovávaným tématům.

Jako významný inspirační zdroj pro charakteristiku současného stavu problematiky citování z vybraných hledisek mi posloužily poznatky a praktické zkušenosti získané z výuky výběrových seminářů na Ústavu informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze. Konkrétně se jednalo o semináře Citování on-line informačních zdrojů (zimní semestr 2007/2008), Citační styly v odborné literatuře (letní semestr 2007/2008) a Citování jako součást přípravy odborného textu (zimní semestry 2008/2009 a 2009/2010). Tato část disertačního projektu byla tedy zpracovávána průběžně po většinu doby doktorského studia.

Paralelně probíhalo zpřesňování vlastního jádra práce. Původně jsem předpokládala, že za účelem splnění cílů práce bude jako první důležitý krok postačovat analýza způsobů zahrnutí citačních dat do existujících XML formátů. V průběhu realizace tohoto kroku jsem však zjistila, že toto omezení by nebylo účelné. Důvod je prostý – v informační praxi se zatím setkáváme s nezanedbatelným počtem formátů nevyužívajících jazyk XML. Tyto formáty přitom v řadě případů patří mezi tradičně a dlouho užívané a představují důležitou součást současné praxe. Zároveň jsem na základě průzkumu vzorku citačních formátů zjistila, že ne vždy je k dispozici jejich jednoznačná a dostatečně podrobná specifikace. Je-li potom k dispozici, ne vždy je pak v praxi dodržována.

Logickým rozšiřujícím krokem se proto staly sběr a analýza dat týkajících se těchto formátů, a to jak ze specifikace formátu, tak z jeho vybraných praktických implementací. Pro tento účel jsem nejdříve vytvořila datové sady z různých informačních systémů, jejichž podobu pak jsem na základě několika kritérií

porovnávala se specifikací formátu. Výsledná zjištění se stala klíčovým podkladem pro formulaci doporučených postupů pro zahrnutí citačních dat do XML formátů.

Protože je disertační projekt zaměřen na XML a obecně strojově srozumitelná data, připravila jsem také stručnou charakteristiku projektů zaměřených na tento typ dat.

Navíc pro automatizované zpracování dat jsou podstatné také patřičné nástroje – proto jsem do disertačního projektu zařadila také oblast citačních manažerů (se zaměřením na používané formáty). Podmínkou pro jejich zahrnutí do projektu se stala jejich dostupnost v podobě otevřeného zdrojového kódu (*open source*) a/nebo dostupnost zdarma prostřednictvím internetového prohlížeče. Tuto podmínku jsem zvolila z praktického důvodu (s nástroji bylo potřeba mít možnost dlouhodobě pracovat bez nutnosti investovat do zakoupení softwaru finanční prostředky) a také v návaznosti na odborný zájem právě o software s otevřeným zdrojovým kódem (mj. o knihovní software Evergreen, ale i o řadu dalších nástrojů využitelných jak v praxi, tak pro výukové účely).

Jako další podstatný podklad pro formulaci doporučených postupů a jejich praktickou aplikaci jsem využila data (záznamy) z různých citačních stylů, resp. z dokumentů, v nichž jsou popsány. Konkrétně byli vybráni tři reprezentanti různých zaměřených citačních stylů – mezinárodní normy ISO⁵ 690 a ISO 690-2 (doplněné o české překlady těchto norem, tj. ČSN ISO 690 a ČSN ISO 690-2), styl ACS⁶ a APA⁷. Zatímco mezinárodní normy nemají určenu specifickou aplikační oblast, styl ACS se samozřejmě zaměřuje na oblast chemie a styl APA na oblast psychologie.

Z výše uvedených norem a citačních stylů jsem excerpovala záznamy dokumentů a část jich zároveň ručně převedla do první pracovní verze XML formátu. Řada poznatků z tohoto prvního pokusného převodu byla následně využita při praktické aplikaci navrženého formátu.

Při práci s citačními daty z různých zdrojů jsem zjistila, že je nezbytné mít k dispozici vhodnou typologii dokumentů. Ve třech zvolených citačních stylech však

⁵ ISO je jako zkratka názvu International Organization for Standardization (*Mezinárodní organizace pro standardizaci*).

⁶ ACS je zkratka názvu *American Chemical Society* (Americká chemická společnost).

⁷ APA je zkratka názvu *American Psychological Association* (Americká psychologická asociace).

byly použity různé typologie dokumentů, které navíc nebyly vždy jednoznačně vzájemně převoditelné.

Z toho důvodu jsem přikročila ke sběru dat o vybraných typologiích dokumentů a jejich následné analýze. Cílem bylo v tomto případě zjistit, které typologie či jaký přístup při jejich tvorbě se jeví jako vhodný s ohledem na zachycení citačních dat v XML formátech.

V tomto případě jsem podobně jako u předchozích analýz nejprve formulovala otázky, na které jsem následně v dostupných zdrojích hledala odpovědi.

Paralelně s výše uvedenými kroky proběhla ještě analýza citační praxe v tzv. barevných knihách IUPAC, tj. aplikační oblasti disertačního projektu.

Hlavní výstupy práce proto představují výsledky jednotlivých analýz a formulace doporučených postupů pro zahrnutí citačních dat do XML formátů spolu s příkladem praktického využití formátu vytvořeného aplikací doporučených postupů. Jako vzorek dat zpracovaných z barevných knih IUPAC jsem zvolila záznamy článků z časopisu Pure and Applied Chemistry, které jsou uvedeny ve třech barevných knihách. I ze zpracování těchto dat jsem získala další podněty pro upřesnění doporučených postupů. Navíc tato data posloužila jako podklad pro vizualizaci citačních vztahů.

Vzhledem ke skutečnosti, že se jedná o velmi rozsáhlou a komplexní oblast, jsem záměrně v jednotlivých částech analyzovala pouze vybrané formáty, styly apod.

2. Současný stav problematiky citování

2.1 Úvod

Citování tvoří integrální součást přípravy vědeckého textu. Autoři uvádějí zdroje, z nichž při přípravě textu vycházeli, na něž navazují, které naopak kritizují apod. Dodržování citační etiky, tj. správné a úplné uvádění použitých zdrojů, je jedním z klíčových principů vědecké práce. Ostatně právě na analýze citačních dat jsou založeny indikátory, na jejichž základě je následně prováděno hodnocení či měření vědy (tj. scientometrie), jednotlivých vědeckých pracovišť, vědeckých pracovníků atd. Využívány jsou různé indikátory, patří mezi ně např. tzv. impakt faktor [47] či h-index [48].

Problematikou citování je možné se zabývat z různých hledisek. V následujících podkapitolách jsou uplatněna některá z nich – nejprve je citování přiblíženo ve vztahu k informační vědě a knihovnicko-informační praxi. Dále je o citování pojednáno jako o součásti informační gramotnosti a následně je nastíněna problematika standardizace způsobů citování – právě nedostatečnou míru standardizace dokládá provedená analýza citačních stylů v nejlépe hodnocených vědeckých časopisech. Na ni navazuje podkapitola věnovaná automatizovanému přebírání citačních dat. Další dvě podkapitoly upozorňují na některé významné změny, ke kterým v oblasti citování dochází v souvislosti s proměnou vztahů mezi aktéry informační společnosti a s rozvojem elektronického publikování.

2.2 *Citování ve vztahu k teorii informační vědy a knihovnicko-informační praxi*

Citování představuje jeden z nejpodstatnějších prostředků, jejichž zkoumáním se zabývá informační věda.⁸ Budeme-li informační vědu charakterizovat jako vědu, která se zabývá komunikací informací ve společnosti,⁹ a budeme-li tedy

⁸ Úzký je pochopitelně rovněž vztah problematiky citování k scientometrii a informetrii a zvláště k bibliometrii – jednou z jejích hlavních metod je citační analýza [50, s. 8].

⁹ Např. A. Merta hovoří o informační vědě, resp. informatice jako o „nauce o vzniku, oběhu a společenském vlivu (působení) informací“ [49, s. 2 elektronické verze textu, s. 10 původní verze textu].

klást důraz nikoliv na informace samotné, ale na jejich komunikaci, přenos (obecně tedy jejich dynamické aspekty), vystupuje role citování ještě více do popředí.

S citováním a souvisejícími otázkami obecně se v knihovnicko-informační praxi lze setkat velmi často, ať se jedná např. o zjišťování citovanosti určitého časopiseckého článku, konkrétního autora, o dohledávání úplného textu dokumentu na základě dostupných bibliografických údajů o něm (převzatých např. z časopiseckého článku, v němž je daný dokument citován) či o vytváření a využívání dokumentů, v nichž jsou citovány použité zdroje. Velmi jednoduše můžeme říci, že citování umožňuje vyjádřit vzájemný vztah mezi různými dokumenty (citující dokument a citovaný dokument).

Z praktických aplikací, v nichž je citování klíčovým jevem, lze zmínit zejména citační rejstříky, ať již produkované ISI¹⁰ a spolu se souvisejícími produkty a službami zveřejňované na platformách Web of Science, resp. Web of Knowledge, firmou Elsevier (jejím produktem je citační databáze Scopus) či například vytvářené automaticky (mj. CiteSeer.IST).

Z terminologického hlediska je třeba poznamenat, že v této práci je používán termín *záznam*¹¹ pro označení českého ekvivalentu anglického výrazu *reference* a termín *citace* pro označení českého ekvivalentu anglického výrazu *citation*, popř. *in-text citation*.

Záznamem je tedy míněn soubor údajů o citovaném zdroji potřebný pro nalezení úplného textu zdroje a typicky uvedený nikoliv v textu, ale v rámci seznamu použitých zdrojů. Můžeme také říci, že se jedná o bibliografický záznam dokumentu vytvořený v souladu s konkrétním citačním stylem (viz kap. 2.4), popř. v této práci i formátu (*record*).

Jako citace je označována stručná zmínka o použitém (citovaném) zdroji, která je součástí textu dokumentu a odkazuje na záznam s podrobnějšími údaji o tomto zdroji. Jinými slovy citace slouží jako odkaz v textu směřující na

¹⁰ ISI je zkratka názvu *Institute for Scientific Information* (Ústav pro vědecké informace).

¹¹ Ačkoliv se v tomto případě nabízí také možnost užití termínu *odkaz*, je dána přednost výrazu „záznam“, a to v souladu s terminologií užitou v příručce E. Bratkové *Metody citování literatury a strukturování bibliografických záznamů podle mezinárodních norem ISO 690 a ISO 690-2 : metodický materiál pro autory vysokoškolských kvalifikačních prací* [51]. Výraz „odkaz“ by navíc bylo možné pokládat za víceznačný právě v souvislosti s digitálním prostředím, kdy jsou jako odkazy často označovány adresy.

bibliografický záznam, jedná se tedy o citaci bibliografického záznamu (a jeho prostřednictvím původního dokumentu) v textu. Tato terminologie se liší od terminologie užitá v platných citačních normách ČSN ISO 690 [52] a ČSN ISO 690-2 [53]. Ty pracují s překlady citace (*reference*) a odkazy (*citations*), které však nevyjadřují přesně podstatu věci a působí spíše matoucím dojmem.

2.3 Citování jako součást informační gramotnosti

Problematice citování je z pochopitelných důvodů věnována pozornost i na Ústavu informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze (viz např. [54], [55], [56] či [57]) a rovněž v knihovnické praxi, zejména ve vysokoškolském prostředí. Citování je totiž jednou z podstatných komponent informační gramotnosti a informačního vzdělávání – praktické návody typu „jak citovat“ jsou často součástí kurzů informační gramotnosti (viz např. [58]). K dispozici jsou také příručky věnované citování, např. [51].

Poznamenejme, že již v roce 1984 se setkáváme s upozorněním na nebezpečí spojená s nejistotou výzkumných pracovníků týkající se správné podoby citací a záznamů a na problematičnost správného určování typů dokumentů v souladu s příručkou pro citování, resp. psaní odborných textů [59]. Citování je však problematické (ba možná i problematičtější) i v současné době, viz např. [60, s. 337], [61] či [62]. V posledních dvou citovaných dokumentech překvapivě zjistíme, že i u studentů knihovní a informační vědy jsou chyby v citování velmi časté – studenti chybují zvláště při citování elektronických dokumentů.

2.4 Standardizace v oblasti citování

Není bez zajímavosti, že zatímco v oblasti katalogizace dokumentů či informačních objektů¹² (tedy v knihovnictví) se standardizace výrazně uplatňuje (mj. používání AACR2¹³, ISBD¹⁴, formátů¹⁵ typu MARC¹⁶ atd.), v oblasti citování, jak již

¹² Jedná se zejména o katalogizaci jmennou, v katalogizaci věcné je situace podstatně složitější (viz např. [63]).

¹³ AACR2 je zkratka názvu *Anglo-American Cataloging Rules, 2nd Revised Edition* (Anglo-americká katalogizační pravidla, 2. revidované vydání).

¹⁴ ISBD je zkratka výrazu *International Standard Bibliographic Description* (mezinárodní standardní bibliografický popis).

¹⁵ Jako formát (*format*) je označována definovaná struktura pro typ informací, které jsou určeny ke zpracování, uložení na magnetická nebo optická média, zobrazení na obrazovce nebo tisku [42].

¹⁶ MARC je zkratka výrazu *Machine Readable Cataloging* (strojem čitelná katalogizace).

bylo naznačeno, tomu tak zatím není (výjimku mohou zčásti tvořit např. formáty typu RIS¹⁷ či BibTeX užívané při komunikaci mezi citačními manažery či informačním systémem a citačním manažerem).

Existuje vysoký počet různých citačních stylů (*citation styles*), tj. sad pravidel pro vytváření citací a záznamů, které určují mj. typ uváděných údajů, jejich pořadí, interpunkci mezi jednotlivými údaji či způsoby typografického zvýraznění určitých údajů v záznamu. Výjimkou nejsou vlastní styly významných časopisů. Některé citační styly jsou používány relativně hojně, například v psychologii se jedná o citační styl vytvářený a aktualizovaný APA. Od autorů časopiseckých příspěvků je typicky vyžadováno zaslání článku se záznamy použitých zdrojů, které jsou uvedeny v daném citačním stylu, nejsou již zasílány původní data v citačních formátech (ta má autor k dispozici v případě, že ke své práci využívá citační manažer nebo obdobný nástroj). Tím je samozřejmě složitější data následně automaticky zpracovat. Tradiční citační formáty (např. RIS či BibTeX) navíc můžeme označit spíše jako formáty, které jsou strojově čitelné (*machine readable*), nikoliv strojově srozumitelné (*machine understandable*).¹⁸

Vyšší míra standardizace v oblasti citačních stylů by nepochybně přispěla i ke snadnějšímu vytváření uživatelsky přívětivých integrovaných informačních systémů, bylo by jednodušší automaticky pracovat se záznamy a propojovat na základě citací a záznamů vlastní texty dokumentů. Navíc je citování koneckonců založeno na předpokladu, že záznamy/citace slouží mj. ke zpětnému vyhledání citovaných dokumentů. Jinými slovy citování prostřednictvím (automatizovaného) přebírání dat z externích zdrojů také podporuje i opakované využití (*reuse*) citovaných zdrojů; záznamy/citace mohou představovat jakýsi pevný bod ve fluidním pojetí kyberprostoru (zakotvení v místě a v čase).

Kromě argumentu týkajícího se snadnějšího vytváření integrovaných informačních systémů lze ve prospěch standardizace uvést i další důvod – stále se rozrůstá počet samostatně „citovatelných“ informačních objektů, často s vlastními identifikátory (a tím i nutnost stále udržovat a rozvíjet citační styly) – někteří

¹⁷ Název formátu RIS vznikl jako zkratka názvu společnosti *Research Information Systems*.

¹⁸ Strojová srozumitelnost hraje podstatnou roli zejména při tvorbě tzv. sémantického webu (viz dále).

poskytovatelé informačních služeb již nabízejí i možnost exportu záznamů dílčích částí dokumentu, např. obrázků, grafů či tabulek (viz např. systémy EBSCOhost¹⁹ [64] či Illustrata [65]). Z hlediska vstupního zpracování, resp. pozdějšího vyhledávání se hovoří o konceptu hloubkového indexování (*deep indexing*), s trochou nadsázky bychom mohli analogicky hovořit o hloubkových citacích (*deep citations*), popř. hloubkových záznamech (*deep references*). Ostatně na rozdíl od záznamů jsou citace vlastně z principu často hloubkové – citovány bývají poznatky či informace uvedené na konkrétních stranách, v konkrétních kapitolách, popř. jiných částech dokumentů.

V různých oborech je navíc potřebné citovat různé speciální typy dokumentů či informačních objektů. Často jsou citovány např. tabulky či datové sady, mnohdy značně různorodým způsobem [66, s. 4]. V řadě případů je také nutné citovat osobní komunikaci, ať již písemně zachycenou či pouze ústní, dále také např. archiválie apod. O těchto případech pokyny pro citování mnohdy mlčí, jejich řešení je tedy na autorovi. Tím je však dále problematizována standardizace v oblasti citování.

S ohledem na výše uvedené skutečnosti jsou nevyhnutelně používány různé metody pro automatickou extrakci citačních dat, např. metoda autonomního indexování citací²⁰ používaná vyhledávačem CiteSeer.IST [67] či metoda flexibilní nekontrolované extrakce citačních metadat²¹ popsaná v [68]. Tento způsob získávání formalizované podoby dat je však často zatížen nepřesnostmi a chybami. Jak poznamenává E. Raymond, kombinace promyšlených datových struktur a průměrného kódu funguje mnohem lépe než jejich obrácená konfigurace [69], [70].

2.5 Citační styly v nejlépe hodnocených vědeckých časopisech

Jak již bylo naznačeno, v praxi se lze setkat s vysokým počtem různých citačních stylů. Často jsou pojmenovány po publikačních manuálech či příručkách. Je to případ již zmiňovaných stylů APA či ACS, dále např. stylu Chicago či AMA²². Významným stylům jako celkům jsou pak věnovány i samostatné webové prezentace (viz např. [71], [72] či [73]).

¹⁹ EBSCO je název společnosti odvozený od jména svého zakladatele (*Elton B Stephens Company*).

²⁰ Pro metodu je používána zkratka ACI (*Autonomous Citation Indexing*).

²¹ Metoda je označována zkratkou FLUX-CiM (*Flexible Unsupervised Extraction of Citation Metadata*).

²² AMA je zkratka názvu *American Medical Association* (Americká lékařská asociace).

V záznamech vytvořených na základě různých citačních stylů se lze setkat s nejednotnou interpunkcí (sloužící k oddělení hodnot jednotlivých dílčích údajů, např. příjmení autora a jeho prvního²³ jména, či jejich bloků, např. autorských údajů od údajů názvových), nejednotným pořadím jednotlivých údajů (celých bloků údajů i jednotlivých údajů v rámci bloků), výjimkou není absence některých běžně uváděných údajů v záznamech (např. v citačním stylu ACS je povoleno neuvádět v záznamech názvy časopiseckých článků).

Ještě ilustrativnější může v tomto případě být např. skutečnost, že z padesáti časopisů s nejvyšší hodnotou impakt faktoru z části Science databáze Journal Citation Reports [74] lze v rámci pokynů pro autory na webu časopisu pouze v osmi případech identifikovat návaznost na širěji užívaný citační styl. Přítomnost samotných pokynů pro autory je již prakticky pravidlem, v této množině časopisů byl identifikován pouze jeden případ, kdy pokyny pro autory dostupné nebyly.

Přehled těchto padesáti časopisů v pořadí podle klesajícího impakt faktoru je uveden v tab. 1. U jednotlivých časopisů je uveden jejich zkrácený název, dále jsou zachyceny identifikátor ISSN²⁴, hodnota impakt faktoru a URL²⁵ pokynů pro autory.

Tab. 1: Pokyny pro autory v časopisech s nejvyšším impakt faktorem

Pořadí	Zkrácený název	ISSN	Impakt faktor	URL pokynů
1.	CA-CANCER J CLIN	0007-9235	74.575	http://www3.interscience.wiley.com/journal/121481132/home/ForAuthors.html
2.	NEW ENGL J MED	0028-4793	50.017	http://authors.nejm.org/help/NewMs.asp#jstyle
3.	ANNU REV IMMUNOL	0732-0582	41.059	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
4.	NAT REV MOL CELL BIO	1471-0072	35.423	http://www.nature.com/nature/authors/gta/index.html
5.	PHYSIOL REV	0031-9333	35.000	http://www.the-aps.org/publications/prv/prvinfo.htm
6.	REV MOD PHYS	0034-6861	33.985	http://rmp.aps.org/files/rmpgui08.pdf

²³ V této práci je upřednostněn výraz „první jméno“ před výrazem „křestní jméno“. Výraz „první jméno“ není spojen se křtem. Na druhou stranu ne ve všech kulturách má první jméno funkci stejnou funkci, proto ani výraz „první jméno“ není za všech okolností zcela optimální.

²⁴ ISSN je zkratka výrazu *International Standard Serial Number* (mezinárodní standardní číslo seriálu).

²⁵ URL je zkratka výrazu *Uniform Resource Locator* (jednotný lokátor zdrojů).

Pořadí	Zkrácený název	ISSN	Impakt faktor	URL pokynů
7.	JAMA-J AM MED ASSOC	0098-7484	31.718	http://jama.ama-assn.org/misc/ifora.dtl
8.	NATURE	0028-0836	31.434	http://www.nature.com/nature/authors/gta/index.html
9.	CELL	0092-8674	31.253	http://www.cell.com/authors
10.	NAT REV CANCER	1474-175X	30.762	http://mts-nrc.nature.com/letters/Author_Guidelines_Reviews_2004.pdf
11.	NAT GENET	1061-4036	30.259	http://www.nature.com/ng/pdf/gta.pdf
12.	ANNU REV BIOCHEM	0066-4154	30.016	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
13.	NAT REV IMMUNOL	1474-1733	30.006	http://mts-nri.nature.com/letters/Author_Guidelines_Reviews_2004.pdf
14.	NAT REV DRUG DISCOV	1474-1776	28.690	http://mts-nrdd.nature.com/letters/Author_Guidelines_Reviews_2004.pdf
15.	LANCET	0140-6736	28.409	http://www.download.thelancet.com/fatcontentassets/authors/artwork-guidelines.pdf
16.	SCIENCE	0036-8075	28.103	http://www.sciencemag.org/about/authors/prep/res/refs.dtl
17.	NAT MED	1078-8956	27.553	http://www.nature.com/nm/pdf/gta.pdf
18.	ANNU REV NEUROSCI	0147-006X	26.405	http://www.annualreviews.org/authors/pdf/green_book_06june06.pdf
19.	NAT REV NEUROSCI	1471-0048	25.940	http://mts-nrn.nature.com/letters/Author_Guidelines_Reviews_2004.pdf
20.	ANNU REV ASTRON ASTR	0066-4146	25.826	http://www.annualreviews.org/authors/pdf/green_book_06june06.pdf
21.	NAT IMMUNOL	1529-2908	25.113	http://www.nature.com/ni/pdf/gta.pdf
22.	NAT PHOTONICS	1749-4885	24.982	http://www.nature.com/nphoton/pdf/nphoton_gta.pdf
23.	CANCER CELL	1535-6108	24.962	http://www.cell.com/cancer-cell/authors
24.	NAT REV GENET	1471-0056	24.185	http://mts-nrg.nature.com/letters/Author_Guidelines_Reviews_2004.pdf
25.	CHEM REV	0009-2665	23.592	http://pubs.acs.org/paragonplus/submit/chreay/chreay_authguide.pdf
26.	NAT MATER	1476-	23.132	http://www.nature.com/nmat/pdf/nma

Pořadí	Zkrácený název	ISSN	Impakt faktor	URL pokynů
		1122		t_gta.pdf
27.	ANNU REV CELL DEV BI	1081-0706	22.731	http://www.annualreviews.org/authors/pdf/green_book_06june06.pdf
28.	NAT BIOTECHNOL	1087-0156	22.297	http://www.nature.com/nbt/pdf/gta.pdf
29.	ANNU REV PLANT BIOL	1040-2519	22.192	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
30.	PHARMACOL REV	0031-6997	21.936	http://pharmrev.aspetjournals.org/site/misc/ifora.xhtml
31.	ANNU REV PHARMACOL	0362-1642	21.561	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
32.	IMMUNITY	1074-7613	20.579	http://www.cell.com/immunity/authors
33.	NAT NANOTECHNOL	1748-3387	20.571	http://www.nature.com/nnano/pdf/nnano_gta.pdf
34.	ANNU REV PHYSIOL	0066-4278	19.740	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
35.	ENDOCR REV	0163-769X	18.562	http://edrv.endojournals.org/misc/itoa.shtml
36.	PHYS REP	0370-1573	18.522	http://www.elsevier.com/wps/find/journaldescription.cws_home/505703/authorinstructions
37.	PROG MATER SCI	0079-6425	18.132	http://www.elsevier.com/wps/find/journaldescription.cws_home/414/authorinstructions
38.	NAT CELL BIOL	1465-7392	17.774	http://www.nature.com/ncb/pdf/ncb_gta.pdf
39.	ANN INTERN MED	0003-4819	17.457	http://www.annals.org/site/misc/ifora.xhtml
40.	CHEM SOC REV	0306-0012	17.419	http://www.rsc.org/Publishing/Resource/AuthorGuidelines/ArticleLayout/sect4.asp
41.	J CLIN ONCOL	0732-183X	17.157	http://jco.ascopubs.org/ifc/prepguide.dtl
42.	ANNU REV BIOPH BIOM	1056-8700	17.049	http://www.annualreviews.org/authors/pdf/blue_handbook_2007.pdf
43.	MICROBIOL MOL BIOL R	1092-2172	16.950	http://mmbr.asm.org/misc/journal-ita_org.dtl
44.	CELL STEM CELL	1934-5909	16.826	http://www.cell.com/cell-stem-cell/authors
45.	NAT PHYS	1745-2473	16.821	http://www.nature.com/nphys/pdf/nphys_gta.pdf
46.	PROG POLYM SCI	0079-6700	16.819	http://www.elsevier.com/wps/find/journaldescription.cws_home/418/author

Pořadí	Zkrácený název	ISSN	Impakt faktor	URL pokynů
				instructions
47.	ALDRICHIM ACTA	0002-5100	16.733	–
48.	B AM MUS NAT HIST	0003-0090	16.692	http://research.amnh.org/scipubs/f/stylemanual.pdf
49.	J CLIN INVEST	0021-9738	16.559	http://www.jci.org/kiosk/publish
50.	CLIN MICROBIOL REV	0893-8512	16.409	http://cmr.asm.org/misc/journal-ita_org.dtl

V pokynech pro autory u těchto 50 nejvýznamnějších časopisů převažuje jako forma pokynů pro citování slovní popis a příklady (43 případů), v šesti případech se pak vyskytoval pouze slovní popis a v jednom případě nebyla data dostupná.

Ve 46 případech byly do pokynů pro autory zařazeny pokyny týkající se záznamů i citací, ve třech případech tomu tak nebylo, v jednom případě nebyla potřebná data k dispozici.

Pouze v osmi případech byla zjištěna návaznost na širěji užívaný citační styl.

Podrobné údaje jsou k dispozici v tab. 2 (údaje týkající se jednotlivých časopisů jsou uspořádány stejně jako v předcházející tabulce).

Tab. 2: Citační styly v pokynech pro autory ve vybraných časopisech

Pořadí	Forma pokynů	Zahrnutí záznamů i citací	Návaznost na citační styly	Název citačního stylu
1.	slovní popis a příklady	ano	ano	AMA Manual of Style (10th ed.)
2.	slovní popis a příklady	ano	ne	–
3.	slovní popis a příklady	ano	ne	–
4.	slovní popis	ano	ne	–
5.	slovní popis a příklady	ano	ne	–
6.	slovní popis a příklady	ano	ne	–
7.	slovní popis a příklady	ano	ano	AMA Manual of Style (10th ed.)

Pořadí	Forma pokynů	Zahrnutí záznamů i citací	Návaznost na citační styly	Název citačního stylu
8.	slovní popis	ano	ne	–
9.	slovní popis a příklady	ne	ne	–
10.	slovní popis a příklady	ano	ne	–
11.	slovní popis	ano	ne	–
12.	slovní popis a příklady	ano	ne	–
13.	slovní popis a příklady	ano	ne	–
14.	slovní popis a příklady	ano	ne	–
15.	slovní popis a příklady	ano	ano	Vancouver style
16.	slovní popis a příklady	ano	ne	–
17.	slovní popis	ano	ne	–
18.	slovní popis a příklady	ano	ne	–
19.	slovní popis a příklady	ano	ne	–
20.	slovní popis a příklady	ano	ne	–
21.	slovní popis	ano	ne	–
22.	slovní popis a příklady	ano	ne	–
23.	slovní popis a příklady	ano	ne	–
24.	slovní popis a příklady	ano	ne	–
25.	slovní popis a příklady	ano	ano	The ACS Style Guide
26.	slovní popis a příklady	ano	ne	–
27.	slovní popis a příklady	ano	ne	–
28.	slovní popis	ano	ne	–
29.	slovní popis a příklady	ano	ne	–
30.	slovní popis a příklady	ano	ne	–
31.	slovní popis a příklady	ano	ne	–

Pořadí	Forma pokynů	Zahrnutí záznamů i citací	Návaznost na citační styly	Název citačního stylu
32.	slovní popis a příklady	ne	ne	–
33.	slovní popis a příklady	ano	ne	–
34.	slovní popis a příklady	ano	ne	–
35.	slovní popis a příklady	ano	ano	CBE Style Manual: A Guide for Authors, Editors and Publishers (5th ed.)
36.	slovní popis a příklady	ano	ne	–
37.	slovní popis a příklady	ano	ne	–
38.	slovní popis	ano	ne	–
39.	slovní popis a příklady	ano	ano	National Library of Medicine
40.	slovní popis a příklady	ano	ne	–
41.	slovní popis a příklady	ano	ne	–
42.	slovní popis a příklady	ano	ne	–
43.	slovní popis a příklady	ano	ne	–
44.	slovní popis a příklady	ne	ne	–
45.	slovní popis a příklady	ano	ne	–
46.	slovní popis a příklady	ano	ne	–
47.	–	–	–	–
48.	slovní popis a příklady	ano	ano	Style Manual Scientific Publications (4th ed.)
49.	slovní popis a příklady	ano	ano	AMA Manual of Style (10th ed.)
50.	slovní popis a příklady	ano	ne	–

Přestože u většiny časopisů nebyla identifikována přímá návaznost na šířeji užívaný citační styl,²⁶ lze konstatovat, že v případech, kdy se jednalo o časopisy vydávané jedním vydavatelem, byly většinou pokyny týkající se citování pro autory jednotlivých vydávaných časopisů podobné nebo shodné. Ze vzorku 50 časopisů bylo možné tuto hypotézu prověřit u 32 časopisů, neboť takový byl celkový počet časopisů, u nichž byl zjištěn výskyt stejného vydavatele jako u alespoň jednoho dalšího časopisu ze vzorku (přehled vydavatelů jednotlivých časopisů je uveden v tab. 3). Konkrétně se jednalo o časopisy vydavatelů Nature Publishing Group (16 výskytů), Annual Reviews (devět výskytů), Cell Press (čtyři výskyty) a Elsevier (tři výskyty).²⁷ U časopisů od vydavatelů Nature Publishing Group, Annual Reviews a Cell Press byla ověřena shodnost nebo výrazná podobnost pokynů, naopak u časopisů vydávaných firmou Elsevier byly zjištěny výrazné rozdíly.

Podrobněji rozebereme časopisy vydávané Nature Publishing Group. V tomto případě byly identifikovány čtyři skupiny velmi podobných znění pokynů pro autory, z nichž ve třech bylo znění pokynů (s výjimkou uvedení názvu daného časopisu) shodné a v jedné velmi podobné:

- skupina časopisů č. 1 a 2 (zcela shodné pokyny),
- skupina časopisů č. 10, 13, 14, 19 a 24 (zcela shodné pokyny, ve všech případech se jednalo o časopisy typu Reviews),
- skupina časopisů č. 11, 17, 21, 28 a 38 (velmi podobné pokyny, liší se zejména v maximálním povoleném počtu záznamů pro jednotlivé typy časopiseckých příspěvků a ve výčtu typů příspěvků, u nichž má nebo naopak nemá v záznamech být uváděn název článků),
- skupina časopisů č. 22, 26, 33 a 45 (zcela shodné pokyny).

²⁶ Například normy ISO 690 a ISO 690-2 nejsou v žádném z uvedených časopisů použity. Za jeden z možných důvodů tohoto stavu lze pokládat i skutečnost, že mezinárodní norma se na rozdíl od publikačních manuálů APA, ACS či dalších věnuje pouze oblasti citování. Nepokrývá tedy ostatní tematické oblasti související s přípravou vědeckých či odborných textů.

²⁷ Názvy vydavatelů byly ověřovány v systémech Journal Citation Reports (součást platformy Web of Knowledge), Ulrichsweb.com a v případě zjištěných diskrepancí rovněž na oficiálních webových prezentacích časopisů. Časopisy s pořadovými čísly 36, 37 a 46 jsou přes rozdílné formy uvádění vydavatele v uvedených zdrojích považovány souhrnně za časopisy vydávané společností Elsevier.

Tab. 3: Vydavatelé vybraných časopisů

Pořadí	Vydavatel
1.	John Wiley & Sons, Inc.
2.	Massachusetts Medical Society
3.	Annual Reviews
4.	Nature Publishing Group
5.	American Physiological Society
6.	American Physical Society
7.	American Medical Association
8.	Nature Publishing Group
9.	Cell Press
10.	Nature Publishing Group
11.	Nature Publishing Group
12.	Annual Reviews
13.	Nature Publishing Group
14.	Nature Publishing Group
15.	The Lancet Publishing Group
16.	American Association for the Advancement of Science
17.	Nature Publishing Group
18.	Annual Reviews
19.	Nature Publishing Group
20.	Annual Reviews
21.	Nature Publishing Group
22.	Nature Publishing Group
23.	Cell Press
24.	Nature Publishing Group
25.	American Chemical Society
26.	Nature Publishing Group
27.	Annual Reviews
28.	Nature Publishing Group
29.	Annual Reviews
30.	American Society for Pharmacology and Experimental Therapeutics
31.	Annual Reviews
32.	Cell Press
33.	Nature Publishing Group
34.	Annual Reviews
35.	The Endocrine Society
36.	Elsevier B.V., North-Holland

Pořadí	Vydavatel
37.	Elsevier B.V., Pergamon
38.	Nature Publishing Group
39.	American College of Physicians
40.	Royal Society of Chemistry
41.	American Society of Clinical Oncology
42.	Annual Reviews
43.	American Society for Microbiology
44.	Cell Press
45.	Nature Publishing Group
46.	Elsevier Ltd.
47.	Aldrich
48.	American Museum of Natural History
49.	American Society for Clinical Investigation
50.	American Society for Microbiology

2.6 Automatizované přebírání citačních dat

Digitální prostředí umožňuje v řadě případů možnost automatizovaného přebírání citačních dat (záznamů). Jedná se o jev typický zejména pro digitální knihovny (např. SpringerLink, ScienceDirect či Wiley InterScience) [5].²⁸ Ty se pak díky pokročilým možnostem automatizované práce se záznamy mohou stát přívětivějšími místy pro uživatele [78].

Podobné možnosti se však pozvolna stávají také součástí nabídky online katalogů knihoven (za všechny jmenujme např. souborný katalog WorldCat). V tomto případě data nabízí přímo jejich producent. Na druhé straně řada citačních manažerů umožňuje automatické převzetí citačních dat z konkrétní webové stránky, i v tomto případě je s ohledem na kvalitu a přesnost převzatých dat výhodou, pokud jejich podobu přizpůsobí producent dat (viz např. možnost přebírat data z katalogu Kongresové knihovny ve Washingtonu a z dalších zdrojů a načítat je do citačního manažeru Zotero).

V tomto ohledu může být inspirativní zmínit také online encyklopedii Wikipedia, která jednak nabízí možnost získání citačních dat týkajících se

²⁸ O citování ve vztahu k digitálním knihovnám je pojednáno též např. v těchto zdrojích: [75], [76] a [77].

konkrétního hesla/článku (viz možnost „Cite this page“ v levém sloupci webové stránky), jednak v samotných heslech/článcích se setkáváme s využitím citačních značek, díky kterým jsou přesněji uváděny jednotlivé prvky záznamů jednotlivých zdrojů (viz [79] a též [80]).

Podstatnou úlohu hraje (a ještě podstatnější roli bude hrát i v budoucnu) také používání identifikátorů – pro digitální prostředí může jako typický příklad posloužit DOI, využívány jsou samozřejmě i další identifikátory, které známe z tradičního prostředí (ISBN²⁹, ISSN apod.), a také identifikátory, s nimiž se setkáváme v e-printových archivech (např. identifikátory v systému arXiv.org) či dalších systémech (např. identifikátory PMID³⁰ v systému PubMed).

2.7 Proměna vztahů mezi aktéry informační sféry společnosti

Oblast citování je charakteristická prolínáním zájmů řady aktérů (autoři, vydavatelé, zprostředkovatelé, uživatelé). Právě v souvislosti s problematikou citování je nezbytné brát v úvahu proměny vzájemných vztahů mezi těmito aktéry, k nimž dochází v souvislosti s rozvojem informačních a komunikačních technologií.

Tyto proměny můžeme v nástinu popsat v návaznosti na teorii informačních systémů. Informační systém v širším slova smyslu zahrnuje (podle [50, s. 4]):

1. systém vytváření informací;
2. systém zprostředkování informací;
3. systém využívání informací.

Vlivem rozvoje internetu a dalších informačních a komunikačních technologií však dochází k prolínání a výraznému vzájemnému ovlivňování všech tří zmíněných systémů (viz také např. [82]). Díky internetu se může dnes každý stát vydavatelem³¹ – role autora a vydavatele se tedy sblíží. Podobně informační systémy používané vydavatelem mohou být například zčásti zpřístupněny uživatelům, čtenářům. Ti zase mohou s využitím informačních a komunikačních technologií velmi snadno poskytovat zpětnou vazbu autorům.

²⁹ ISBN je zkratka výrazu *International Standard Book Number* (mezinárodní standardní číslo knih).

³⁰ PMID je zkratka výrazu *PubMed Identifier* (identifikátor PubMed) nebo *PubMed unique Identifier* (unikátní identifikátor PubMed).

³¹ Vydavatelem se může stát a stává také knihovna, názorně to doložily např. příspěvky na semináři CASLIN 2007 (viz blíže podrobná zpráva ze semináře [81]).

V důsledku výše charakterizovaných změn se i z autora a zároveň vydavatele může snadno stát „katalogizátor“ – svá díla, která si sám vydá, si může i sám popsat, například prostřednictvím metadat³² v hlavičce HTML.^{33,34} Tím však publikační řetězec nekončí – ten, kdo si dílo přečte, jej může chtít citovat ve svém vlastním díle.

Lze však konstatovat, že popisem zdrojů (nejedná-li se o stručný popis webových zdrojů prostřednictvím metadat apod.) se stále zabývají především knihovníci jako odborníci na katalogizaci. U citování je tomu však právě naopak – citují především odborníci a vědci, kteří jsou v oblasti citování (či šířeji zpracování informací) laiky. Přesto se zdá, že některé citační styly jsou velmi složité na pochopení (ať již z hlediska užívané terminologie či kupříkladu typologie dokumentů), a to i pro odborníky v oblasti zpracování informací.

2.8 Elektronické publikování jako změna paradigmatu publikování

Samotné elektronické publikování lze považovat za změnu paradigmatu publikování, a to podobně výraznou, jako znamenal vynález knihtisku (o důsledcích vynálezu knihtisku podrobně pojednává mj. B. Anderson [84]). Přitom nyní se stále nacházíme v období počátků elektronického publikování – na řadu podstatných otázek souvisejících s elektronickým publikováním zatím nebyly nalezeny uspokojivé odpovědi (byly-li už vůbec takové otázky formulovány). Rozvineme-li paralelu se vznikem knihtisku, můžeme prohlásit, že žijeme ve věku digitálních „prvotisků“ (*digital incunabula*) [85]. Málodky si dnes uvědomujeme např. skutečnost, že v počátcích knihtisku nebylo samozřejmé, že všechny exempláře dané vytištěné knihy byly totožné (viz např. [86, s. 33]).

V oblasti elektronického publikování čelíme obdobným problémům, kdy je např. třeba počítat s tím, že e-print zveřejněný v online archivu či obecně na internetu může, ale nemusí přesně odpovídat později zveřejněnému článku. Jedná-li se skutečně o preprint (a nikoliv o postprint či o přesnou verzi již publikovaného časopiseckého článku), pak v něm ještě nebývají reflektovány připomínky

³² Metadata jsou strojově srozumitelné informace o webových zdrojích nebo dalších věcech [43].

³³ HTML je zkratka výrazu *HyperText Markup Language* (značkovací jazyk pro hypertext).

³⁴ Existují rovněž mj. různé snahy o standardizace v oblasti citování, např. zavedení volitelných tagů do jazyka HTML, které by uvozovaly bibliografické údaje [83].

recenzentů. Rovněž případy, kdy je jeden dokument dostupný ve více formátech či na více místech na internetu, mohou být z hlediska určení stejnosti či jinakosti „exemplářů“ problematické a těžko řešitelné.

V souvislosti s tím lze upozornit na další problematickou skutečnost – v řadě případů je v praxi citována tzv. úvodní stránka (*splash page*),³⁵ která je určena pro uživatele-člověka a slouží jako rozcestník k různým verzím dokumentu (ať již v různých formátech, např. HTML či PDF³⁶, nebo skutečně různým verzím dokumentu, jimiž jsou např. preprint či postprint). V případě odlišnosti různých verzí z hlediska obsahu či uspořádání může být odkaz na URL této úvodní stránky problematický.

Ilustrativní je také hledání způsobů, jak citovat konkrétní části elektronických knih. Například čtečka elektronických knih Amazon Kindle k tomu používá tzv. lokace (*locations*), tj. úseky o pevné délce. Její konkrétní hodnota je uváděna pouze přibližně (cca 128 bytů), neboť není součástí oficiální dokumentace (jedná se pouze o názor objevující se v internetových diskusních fórech, viz např. [89]). Lokacemi je označována nejenom vlastní textová část elektronické knihy, ale lokace jsou přiděleny také prelimináriím a jakýmkoli případným dalším částem knihy. Výhodou konceptu lokací je skutečnost, že umožňují přesnější uvádění místa v textu než klasické stránky. Je to také pravděpodobně jediný vhodný způsob citování konkrétního místa v textu elektronické knihy zobrazené v této čtečce – uživatel si totiž může text (je-li ve formátu, který to umožňuje) podle potřeby zvětšit či zmenšit. Počet „stránek“ se tudíž v návaznosti na to může zvětšovat či zmenšovat. Na druhou stranu převoditelnost lokací na původní stránky v „klasické“ podobě knihy nemusí být jednoduchou záležitostí.

Další podobné znaky počátků elektronického publikování a počátků knihtisku jsou nasnadě. První inkunábule se záměrně svým vzhledem prakticky nelišily od rukopisů [90, s. 77]. Řada elektronických dokumentů má nyní podobu, která víceméně odpovídá původní podobě dokumentu v tištěné podobě. Příkladem

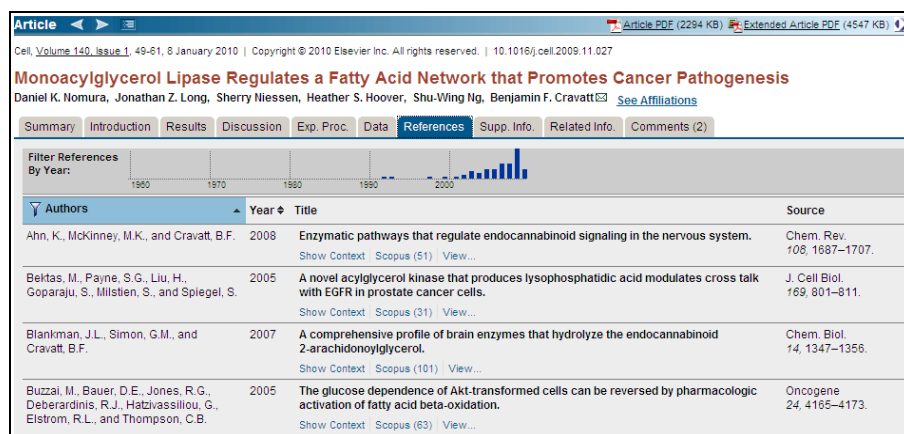
³⁵ Toto označení je užíváno ve výstupech z projektu Iniciativa otevřených archivů – znovuvyužití a výměna objektů (*Open Archives Initiative – Object Reuse and Exchange*, OAI-ORE) zaměřeného na agregaci webových zdrojů [87], [88].

³⁶ PDF je zkratka výrazu *Portable Document Format* (přenosný formát dokumentů).

může být časopisecký článek zveřejněný ve formátu PDF původně určeném primárně pro tisk.³⁷

Postupně se již vyvíjejí různé progresivní formy elektronického publikování. Např. D. Shotton hovoří o tzv. sémantickém publikování (*semantic publishing*), které definuje jako „cokoliv, co obohacuje význam publikovaného časopiseckého článku, usnadňuje jeho automatizované nalezení, umožňuje jeho propojení na sémanticky příbuzné články, poskytuje přístup k datům v článku ve využitelné podobě nebo usnadňuje integraci dat mezi příspěvky“ [92, s. 86]. Tentýž autor zdůrazňuje potřebu zpřístupňovat údaje o článku ve strojově čitelné podobě³⁸ a tímtež způsobem zpřístupňovat i seznamy použitých zdrojů tak, aby bylo možné vytvářet citační sítě [92, s. 91].

Podnětný příklad „článků budoucnosti“ nabízí od počátku roku 2010 časopis Cell. Články jsou strukturovány do částí, které si lze samostatně zobrazovat, záznamy použité literatury jsou pak obohaceny například o možnost získání úplného textu a zjištění citací daného (v tomto případě také citovaného) článku v citační databázi Scopus (viz obr. 1).



Authors	Year	Title	Source
Ahn, K., McKinney, M.K., and Cravatt, B.F.	2008	Enzymatic pathways that regulate endocannabinoid signaling in the nervous system.	Chem. Rev. 108, 1687–1707.
Bektas, M., Payne, S.G., Liu, H., Goparaju, S., Milstien, S., and Spiegel, S.	2005	A novel acylglycerol kinase that produces lysophosphatidic acid modulates cross talk with EGFR in prostate cancer cells.	J. Cell Biol. 169, 801–811.
Blankman, J.L., Simon, G.M., and Cravatt, B.F.	2007	A comprehensive profile of brain enzymes that hydrolyze the endocannabinoid 2-arachidonoylglycerol.	Chem. Biol. 14, 1347–1356.
Buzzai, M., Bauer, D.E., Jones, R.G., Deberardinis, R.J., Hatzivassiliou, G., Elstrom, R.L., and Thompson, C.B.	2005	The glucose dependence of Akt-transformed cells can be reversed by pharmacologic activation of fatty acid beta-oxidation.	Oncogene 24, 4165–4173.

Obr. 1: Ukázka části obohaceného seznamu použitých zdrojů v článku budoucnosti z časopisu Cell

Objevuje se také potřeba jednoznačně identifikovat a citovat nejenom celé dokumenty, ale také jejich dílčí části (např. věty, odstavce, obrázky apod.). Jak již bylo naznačeno, v oblasti elektronických knih (Amazon Kindle) se setkáváme s tzv.

³⁷ Stručný a výstižný přehled rozdílů mezi vybranými formáty používanými v oblasti elektronického publikování a jejich výhod a nevýhod (konkrétně HTML, SGML, PDF a XML) podává [91].

³⁸ Z kontextu je však zřejmé, že je míněna spíše strojově srozumitelná podoba.

lokacemi, které mohou být pro tento účel využity. Lokace se však vztahují vždy ke konkrétnímu elektronickému dokumentu, neslouží jako samostatný identifikátor. Naproti tomu identifikátor DOI³⁹ může figurovat i samostatně. V praxi je již používán právě i pro dílčí části dokumentů. Jako příklad lze uvést hesla ze Zlaté knihy (viz podrobněji kap. 7) či obrázky a tabulky v článcích zveřejněných v časopise PLoS ONE.

Vyskytují se také návrhy citování pouze prostřednictvím identifikátorů. V tomto případě je však nezbytné, aby se jednalo výhradně o identifikátory trvalé (perzistentní). Zaručení skutečné trvalosti však není jednoduché. Navíc řada online zdrojů nemá přidělen žádný trvalý identifikátor. Tradiční URL tuto roli neplní, neboť jednotlivé stránky či jiné dokumenty mohou být v průběhu času přemíst'ovány jinam (na jiná URL) nebo mohou být zcela odstraněny. Problémem mizení obsahu z webu byla a je věnována řada výzkumů (jako příklad lze uvést [93]).

Vrátíme-li se ještě k počátkům knihtisku, můžeme doplnit, že určitou paralelu lze spatřovat i ve skutečnosti, že v okamžiku, kdy díky snadnému rozšiřování kopií knih přestávaly knihy být vzácnými, vzrůstala i obliba přenositelného formátu (*portable format*). Lidé si totiž zvykli s sebou brát knihy všude, aby je mohli kdykoliv využívat [90, s. 88]. V dnešní době jsou podobnými přenositelnými formáty např. formáty určené pro elektronické knihy (např. EPUB⁴⁰ či Mobipocket PRC⁴¹). Jedná se o formáty určené pro čtečky elektronických knih, tj. nejenom pro klasické stolní počítače, ale zejména pro mobilní zařízení, např. právě hardwarové čtečky elektronických knih či chytré mobilní telefony.

2.9 Dílčí závěry

Citování představuje jednu z důležitých oblastí informační vědy i knihovnicko-informační praxe. Na citování je založena existence citačních rejstříků, které jsou využívány mj. při hodnocení vědy, jako zvláštní téma je vymezována například citační etika. Navíc problematika citování tvoří nedílnou součást informační gramotnosti či obecně práce s vědeckými či odbornými informacemi.

³⁹ DOI je zkratka výrazu *Digital Object Identifier* (identifikátor digitálních objektů).

⁴⁰ EPUB je zkratka výrazu *Electronic Publication*, tj. elektronická publikace.

⁴¹ PRC je zkratka výrazu *Palm Resource Compiler*, tj. kompilátor zdrojů pro (operační systém) Palm.

Přesto je zatím standardizace v této oblasti problematická. Skutečnost, že v oblasti vědecké literatury nejsou uplatňovány jednotné postupy pro tvorbu záznamů citovaných dokumentů, názorně doložil průzkum citačních stylů používaných v padesáti časopisech s nejvyšší hodnotou impakt faktoru (za rok 2008) v databázi Journal Citation Reports (Science Edition), která je dostupná v rámci platformy Web of Knowledge. Z těchto padesáti časopisů byl pouze u osmi z nich v pokynech pro autory zmíněn šířeji používaný citační styl. U časopisů vydávaných stejným vydavatelem pak bylo většinou možné setkat se s podobnými nebo shodnými pokyny týkajícími se citování (např. u 16 časopisů z uvedeného vzorku se jednalo o vydavatele Nature Publishing Group). Existuje sice mezinárodní norma zaměřená na oblast citování (ISO 690 a její druhá část ISO 690-2), ta je však v praxi užívána spíše výjimečně. Ze zmíněných padesáti časopisů s nejvyšším impakt faktorem není v žádném z nich aplikována.

V záznamech vytvořených na základě různých citačních stylů se lze setkat s nejednotnou interpunkcí (sloužící k oddělení hodnot jednotlivých dílčích údajů, např. částí jména, či jejich bloků, např. autorských údajů od údajů názvových), nejednotným pořadím jednotlivých údajů (celých bloků údajů i jednotlivých údajů v rámci bloků), výjimkou není absence některých běžně uváděných údajů v záznamech (např. v citačním stylu ACS je povoleno neuvádět v záznamech názvy časopiseckých článků).

V souvislosti s nedostatečnou standardizací začaly vznikat návrhy řešení spočívající v automatické extrakci citačních dat. Tento způsob získávání formalizované podoby dat je však často zatížen nepřesnostmi a chybami.

V současné době navíc dochází k nárůstu významu elektronického publikování, byť zatím na řadu otázek s ním souvisejících nebyla nalezena uspokojivá odpověď. Řadu dat, mj. právě citačních, je však již dnes možné přebírat automaticky, a to jak z prostředí digitálních knihoven, tak i např. z některých knihovních katalogů.

3. Strojům srozumitelná citační data

3.1 Úvod

V souvislosti s rozvojem aplikací sémantického webu⁴² jsou realizovány projekty směřující k vylepšení způsobů zachycení citačních dat ve strojově srozumitelné podobě. Existují různé možnosti, jak tato vylepšení prakticky realizovat – mj. prostřednictvím dedikovaných jazyků či formátů nebo prostřednictvím tzv. ontologií (*ontologies*), tj. podle dnes již klasické definice formulované T. Gruberem formálních, explicitních specifikací (sdílených) konceptualizací [44, s. 199].^{43, 44}

Kromě toho k zachycení citačních dat ve strojově srozumitelné podobě lze využít také XML formáty, které nejsou primárně zaměřeny na citační data, ale je možné je jejich prostřednictvím také vyjádřit, ať již více či méně sofistikovaně.

Podrobněji je o jednotlivých projektech a formátech pojednáno v následujících podkapitolách.

3.2 Vybrané projekty

V této podkapitole je věnována pozornost čtyřem projektům z oblasti práce se strojově srozumitelnými citačními daty, které lze pokládat za inspirativní a přínosné. Jedná se o projekt jazyka pro citační styly, dva projekty ontologií z oblasti citování, resp. práce s bibliografickými daty a o projekt s cílem zjednodušit práci s citačními daty pro uživatele kancelářského balíku OpenOffice.org.

3.2.1 CSL

CSL⁴⁵ je otevřený XML formát určený pro formátování citací a záznamů podle různých citačních stylů. Formát CSL začal jeho autor Bruce D'Arcus původně vyvíjet pro potřeby kancelářského balíku OpenOffice (resp. OpenOffice.org). Verze 0.8 jazyka byla zveřejněna v roce 2009, od března 2010 je aktuální verze 1.0 [95].

⁴² Jako sémantický web (*semantic web*) je označován web dat s významem, jemuž může počítačový program dostatečně porozumět, zjistit, co data znamenají, a podle toho je zpracovat [46].

⁴³ Definice je v tomto kontextu uvedena v upravené podobě – úprava spočívala v převedení definice do plurálu a v jejím obohacení o adjektivum „sdílených“. To totiž přímo součástí Gruberovy definice není, z pozdějších interpretací a také z kontextu Gruberova příspěvku je však zřejmé, že se jedná o implicitní součást definice.

⁴⁴ O významu ontologií ve vztahu k sémantickému webu je pojednáno mj. v [94].

⁴⁵ CSL je zkratka názvu *Citation Style Language* (jazyk pro citační styly).

Původně byla webová prezentace jazyka umístěna na platformě SourceForge.net [96], nově je k dispozici zcela samostatná webová prezentace [97], přičemž vlastní dokumentace jazyka je uložena v repozitáři CSL Docs provozovaném na platformě Bitbucket [98].

Jazyk CSL je nyní využíván především v citačních manažerech Zotero [99] a Mendeley [100]. Právě na webu nástroje Zotero je k dispozici speciální stránka, odkud si mohou uživatelé do svých instalací tohoto zásuvného modulu přidávat jednotlivé styly, zároveň vývojáři (tvůrci) nových stylů mohou tyto styly sdílet prostřednictvím systému SVN/Trac [101].

Na webu manažeru Zotero je rovněž k dispozici přehled syntaxe jazyka CSL, resp. jeho verze 0.8 [102]. Vlastní specifikace jazyka CSL (včetně jeho nejnovější verze a aktualizace) je dostupná prostřednictvím oficiálního webu jazyka [103].

Ve specifikaci verze 1.0 formátu CSL je doporučeno, aby každý styl napsaný v tomto formátu začínal XML deklarací s upřesněním verze jazyka XML a kódování. Kořenovým prvkem ve formátu CSL je pak prvek „cs:style“. Tomu může být přímo podřazeno celkem pět prvků: „cs:info“, „cs:citation“, „cs:bibliography“, „cs:macro“ a „cs:locale“. Zatímco první dva z těchto prvků jsou pro každý styl povinné, tři následující jsou volitelné. Prvek „cs:info“ slouží k uvedení metadat týkajících se stylu jako celku, prvek „cs:citation“ slouží k popisu instrukcí pro formátování citací, prvek „cs:bibliography“ pro formátování celých záznamů, resp. jejich seznamu, prvek „cs:macro“ je určen pro znovuvyužití formátovacích instrukcí (v rámci jednoho stylu i různých stylů), prvek „cs:locale“ pak k překonání implicitní lokalizace [104]. Formát CSL také mj. umožňuje vytváření tzv. závislých stylů (*dependent styles*), jejichž využití je výhodné zejména v případě, kdy více časopisů používá stejný způsob citování.

Příklad části stylu 6. vydání publikačního manuálu APA je k dispozici na obr. 2.

```

<bibliography>
<option name="hanging-indent" value="true"/>
<option name="et-al-min" value="8"/>
<option name="et-al-use-first" value="7"/>
<option name="entry-spacing" value="0"/>
<option name="line-spacing" value="2"/>
<sort>
<key macro="author"/>
<key macro="issued-year" sort="ascending"/>
</sort>
<layout>
<group suffix=".">
<text macro="author" suffix="."/>
<text macro="issued" suffix=" "/>
<group delimiter=".">
<text macro="title"/>
<group>
<text macro="container-contributors"/>
<text macro="secondary-contributors"/>
<group delimiter=",">
<text variable="container-title" font-style="italic"/>
<text variable="collection-title"/>
</group>
</group>
</group>
</group>
<text macro="locators"/>
<group delimiter="," prefix=" ">
<text macro="event"/>
<text macro="publisher"/>
</group>
</group>
<text macro="access" prefix="."/>
</layout>
</bibliography>
</style>

```

Obr. 2: Část stylu APA (6th ed.) v CSL

Jako zajímavost je možné doplnit, že v repozitáři stylů na webu citačního manažeru Zotero je k dispozici i styl „ISO-690 (Author-Date format)“, byť je zatím označen jako nedokončený [101]. Jiná verze stylu určeného pro formátování citací a záznamů v souladu s mezinárodními normami ISO 690 a 690-2 byla vytvořena na Vysokém učení technickém v Brně [105].

3.2.2 BIBO

BIBO⁴⁶ představuje jiný typ projektu než CSL – zatímco formát CSL se zaměřuje na transformaci dat tak, aby byla prezentována v souladu s určitými pravidly (citačními styly), BIBO je ontologií. Na pracích spojených s vývojem ontologie BIBO se podílel mj. Bruce D'Arcus, jehož zásluhou byl vytvořen formát CSL, jak již bylo zmíněno výše.

⁴⁶ BIBO je zkratka názvu *Bibliographic Ontology* (Bibliografická ontologie).

Ontologie BIBO je zaměřena na oblast konceptů (pojmu) a vlastností pro popis citací a záznamů, resp. jejich reprezentace v prostředí sémantického webu. Je založena na využití standardů W3C⁴⁷, zejména XML, XML Namespaces, RDF⁴⁸ a OWL⁴⁹. Pro tuto ontologii je charakteristická rozšiřitelnost – právě díky využití RDF je možné konzistentním způsobem pracovat s různými slovníky. Aktuální verze ontologie je označena jako revize 1.3 – byla zveřejněna v listopadu 2009 [106], [107].

Názorný přehled o ontologii jako celku (včetně počtu tříd, vlastností objektů i dat a dále individuů, tj. instancí) i jejích jednotlivých částech poskytuje vizualizace vytvořená prostřednictvím prohlížeče ontologií [108] a dostupná na platformě Google Code (viz obr. 3).

The image shows a screenshot of a web-based ontology visualization tool. It is divided into several panels:

- Contents:** A list of navigation links:
 - bibo
 - All Resources
 - All Classes (69)
 - All Object Properties (52)
 - All Data Properties (54)
 - All Individuals (14)
- All Resources (189):** A scrollable list of classes and properties:
 - owl:Thing
 - bibo:abstract
 - bibo:AcademicArticle
 - bibo:affirmedBy
 - bibo:annotates
 - bibo:argued
 - bibo:Article
 - bibo:asin
 - bibo:AudioDocument
 - bibo:AudioVisualDocument
 - bibo:authorList
 - bibo:bdarcus
 - bibo:Bill
- Asserted Class Hierarchy:** Shows the inheritance structure:
 - + owl:Thing
 - + bibo:Document ≡ foaf:Document
 - + **bibo:Article**
 - + foaf:Document ≡ bibo:Document
 - + **bibo:Article**
 - bibo:AcademicArticle
- Annotations (4):** Lists specific annotations for the selected class:
 - isDefinedBy: "http://purl.org/ontology/bibo/" (anyURI)
 - term_status: "stable"
 - comment: "A written composition in prose, usually nonfiction, on a specific topic, forming an independent part of a book or other publication, as a newspaper or magazine." (en)
 - label: "Article" (en)
- Superclasses (1):** Lists the superclass:
 - bibo:Document
- Usage (1):** Shows where the class is used:
 - bibo:Issue ≡ dcterms:hasPart only **bibo:Article**

Buttons for "OWL HTML inside" are visible at the bottom of the "Contents" and "Usage" panels.

Obr. 3: Část vizualizace ontologie BIBO

⁴⁷ W3C je zkratka názvu *World Wide Web Consortium*.

⁴⁸ RDF je zkratka výrazu *Resource Description Framework* (Rámec pro popis zdrojů).

⁴⁹ OWL znamená *Web Ontology Language* (jazyk pro ontologie ve webovém prostředí).

Příklady praktického uplatnění ontologie jsou pak k dispozici přímo na oficiálním webu ontologie [109]. Jedním z příkladů je záznam časopiseckého článku (viz obr. 6).

Ontologie BIBO je již implementována v řadě projektů – jejich přehled je k dispozici v rámci webové prezentace ontologie [110]. Jedním z těchto projektů je projekt nazvaný „Zotero BIBO Converter“, jehož cílem je umožnit převod citačních dat z citačního manažeru Zotero do RDF, dalším z projektů je LCCN Permalink Service, tj. služba, která umožňuje využití trvalých URL pro bibliografické záznamy s identifikátory LCCN⁵⁰.

Ontologie je dostupná pod licencí Creative Commons Attribution 1.0 Generic. Za podmínky uvedení autora je tedy ontologii možné volně šířit a upravovat [107], [111].

3.2.3 CiTO

CiTO⁵¹ je název ontologie, která se na rozdíl od výše charakterizované ontologie BIBO zaměřuje na přesnější vyjádření vztahu mezi citovanými a citujícími díly a na vystižení záměru citujícího autora. Ontologie tak reaguje na skutečnost, že ne vždy autoři citují jiná díla ze stejného důvodu. Umožňuje také zachycení frekvence citací, resp. rozlišení případu celkového počtu citací a lokálního počtu citací v textu odkazujících na jiný dokument.

Hlavním hybatelem vývoje ontologie CiTO je David M. Shotton, vedoucí Výzkumné skupiny pro bioinformatiku (*Bioinformatics Research Group*) působící na Katedře zoologie Univerzity v Oxfordu (*Department of Zoology, University of Oxford*).

Pokud se týká popisu citovaných děl, ontologie vychází z principů FRBR⁵² [112]. Typy dokumentů pak vycházejí převážně ze slovníku Dublinského jádra⁵³ [113] a ze slovníku používaného v ontologii SWAN⁵⁴ [114]. [115], [8], [31].

⁵⁰ LCCN je zkratka výrazu *Library of Congress Control Number* (kontrolní číslo Kongresové knihovny).

⁵¹ CiTO je zkratka výrazu *Citation Typing Ontology* (Ontologie typů citací).

⁵² FRBR je zkratka výrazu *Functional Requirements for Bibliographic Records* (Funkční požadavky na bibliografické záznamy).

⁵³ Pro Dublinské jádro (*Dublin Core*) je v praxi často užívána i zkratka DC.

⁵⁴ SWAN je zkratka z názvu *Semantic Web Applications in Neuromedicine* (Aplikace sémantického webu v neuromedicíně).

Ontologie CiTO je napsána v jazyce OWL. Aktuální verze ontologie CiTO je verze 1.6, která byla zveřejněna v roce 2010. Ukázka z této verze ontologie je ve vizualizované podobě k dispozici na obr. 4, který také podává rámcový přehled o rozsahu ontologie (počet tříd, vlastností objektů a dat).

Na obr. 4 jsou rovněž zachyceny některé vztahy, které mohou konkretizovat konkrétní způsob užití citovaného díla.⁵⁵ Jedná se např. o vztahy, resp. vlastnosti „cito:confirms“ (potvrzení), „cito:credits“ (uvedení zdroje), „cito:disagreesWith“ (vyjádření nesouhlasu) či „cito:extends“ (rozšíření).

Obr. 4: Část vizualizace ontologie CiTO

Ukázka uplatnění vztahů definovaných v ontologii CiTO je zařazena i v preprintu příspěvku *CiTO, the Citation Typing Ontology* D. Shottona, kde je ontologie využita k zachycení vztahu tohoto preprintu (resp. citací v něm uvedených)

⁵⁵ Na skutečnost, že autory k citování vedou různé důvody, upozorňují i další autoři, mj. R. G. McInnis a D. Symes [116] či E. Garfield [117, s. 72-73].

k citovaným dokumentům [8]. Existující vztahy jsou v textové podobě zachyceny i u záznamů použité literatury, jak ilustruje obr. 5.

References	
1.	Reis RB, Ribeiro GS, Felzemburgh RDM, Santana FS, Mohr S, Melendez AXTO, Queiroz A, Santos AC, Ravines RR, Tassinari WS, Carvalho MS, Reis MG, Ko AI: Impact of environment and social gradient on <i>Leptospira</i> infection in urban slums. <i>PLoS Neglected Tropical Diseases</i> 2008, 2 : e228. doi: 10.1371/journal.pntd.0000228 CiTO: <i>cites for information, uses data from, Research Paper, Journal Article, peer reviewed.</i>
2.	Shotton D, Portwin K, Klyne G, Miles A: Adventures in semantic publishing: exemplar semantic enhancement of a research article. <i>PLoS Computational Biology</i> 2009 5 : e1000361. doi: 10.1371/journal.pcbi.1000361 . CiTO: <i>cites for information, shares authors with, Research Paper, Journal Article, peer reviewed.</i>

Obr. 5: Zachycení citačních vztahů v preprintu *CiTO, the Citation Typing Ontology* D. Shottona

Ontologie je dostupná pod licencí Creative Commons Attribution 2.5 Generic [7], [118], tj. ontologii je možné volně šířit a upravovat, vždy však s uvedením původního autora.

3.2.4 OOoBib

OOoBib⁵⁶ je název projektu, jehož počátky sahají již do roku 2003 [119]. Zakladatelem projektu se stal David N. Wilson, projekt vede ve spolupráci s Brucem D'Arcusem, jehož jméno bylo zmíněno již v souvislosti s formátem CSL a ontologií BIBO.

Cílem projektu OOoBib je zjednodušit ukládání a správu citačních dat, formátování dokumentů a spolupráci s ostatními uživateli. Projekt by měl přispět k vylepšení souvisejících funkcí kancelářského balíku OpenOffice.org, resp. textového editoru OpenOffice.org Writer. Ten by následně měl uživatelům poskytovat podobný komfort jako komerční citační softwarové nástroje [120].

V souvislosti s projektem OOoBib je rozvíjen procesor CiteProc ve verzích pro různé programovací jazyky (Javascript, Haskell, PHP⁵⁷, Python a Ruby). Procesor

⁵⁶ OOoBib je zkratka názvu *OpenOffice Bibliographic Project* (Bibliografický projekt OpenOffice).

⁵⁷ PHP je rekurzivní zkratka výrazu *PHP: Hypertext Preprocessor* (PHP: Hypertextový preprocesor).

je založen na využití formátu CSL, ale je širěji použitelný (není určen pouze pro citační manažer Zotero) [121], [122], [123].

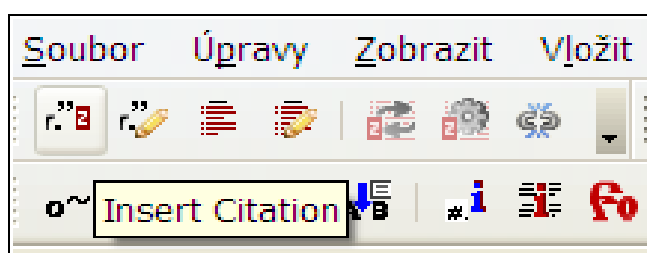
```
<info:doi/10.1134/S0003683806040089> a bibo:Article ;
  dc:title "Effect of argillaceous minerals on the growth of phosphate-mobilizing bacteria Bacillus subtilis"@en ;
  dc:date "2006-01-01" ;
  dc:isPartOf <urn:issn:23346587> ;
  bibo:volume "42" ;
  bibo:issue "4" ;
  bibo:pageStart "388" ;
  bibo:pageEnd "391" ;
  dc:creator <http://examples.net/contributors/2> ;
  dc:creator <http://examples.net/contributors/1> ;
  bibo:authorList ( <http://examples.net/contributors/2> <http://examples.net/contributors/1> ) .

<urn:issn:23346587> a bibo:Journal ;
  dc:title "Applied Biochemistry and Microbiology"@en ;
  bibo:shortTitle "App Biochem and Biol"@en .
```

Obr. 6: Záznam časopiseckého článku v ontologii BIBO

Vzhledem ke skutečnosti, že hlavní stránka projektu [120] i hlavní stránka wiki pro vývojáře [124] byla naposledy aktualizována v listopadu 2008, je však nutné konstatovat, že tento projekt není (s výjimkou procesoru CiteProc) v poslední době příliš rozvíjen.

Pozornost je zatím věnována spíše vylepšování způsobů spolupráce kancelářského balíku s jinými citačními manažery, zejména se Zoterem. Právě tento software je nejvýrazněji doporučován. Uživatel Zotera si může nainstalovat zásuvný modul, který mu umožní s daty ze Zotera pracovat přímo v prostředí textového editoru OpenOffice.org Writer (viz obr. 7).



Obr. 7: Panel zásuvného modulu Zotera v textovém editoru OpenOffice.org Writer

3.3 Citační data ve vybraných XML formátech

Vzhledem ke skutečnosti, že v oblasti STM existuje vysoký počet různých XML formátů používaných pro různé účely, bylo v rámci disertačního projektu

vybráno k analýze z hlediska způsobu zachycení citačních dat pouze několik reprezentativních formátů.

Konkrétně se jedná o sedm formátů – o formáty DocBook, OpenDocument Format for Office Applications, EPUB, TEI, CML, ThermoML a interní XML formát využitý při přípravě XML verze Zlaté knihy. Zatímco první čtyři zmíněné formáty lze zjednodušeně charakterizovat jako formáty určené pro relativně široké spektrum aplikačních oblastí (byť např. formát DocBook je primárně určen pro dokumenty z počítačové oblasti), u zbývajících tří formátů je předpokládáno především využití v oblasti chemie. Formáty ThermoML a interní XML formát Zlaté knihy jsou navíc přímo spjaty s organizací IUPAC, jejíž tzv. barevné knihy představují aplikační oblast disertace.⁵⁸

V následujícím textu jsou u jednotlivých XML formátů zachyceny odpovědi na následující otázky (bylo-li možné je z dostupných zdrojů získat):

- Které prvky a atributy je možné použít k zachycení citačních dat?
- Jak je definován způsob práce s těmito prvky a atributy?
- Jsou u atributů využity slovníky hodnot?
- Jakým způsobem jsou citační data strukturována?
- Využívá formát hierarchie jednotlivých prvků s citačními daty?

Pozornost je přitom zaměřena pouze na oblast zachycení dat o externích informačních zdrojích, tj. nikoliv na možnosti vzájemného odkazování jednotlivých částí dokumentu, popř. na uvádění metadat týkajících se daného dokumentu (souboru).

Zároveň je u každého formátu uvedeny základní údaje o formátu jako takovém (doba vzniku, nejnovější verze, dostupnost jeho specifikace), pokud jsou k dispozici.

3.3.1 DocBook

Počátky vývoje formátu DocBook – mimochodem původně určeného pro práci s dokumentací týkající se operačního systému Unix – sahají do roku 1991.

⁵⁸ Podrobnější zhodnocení využití XML formátů v chemii je k dispozici v [125].

Návrh a první implementace formátu byly realizovány pod hlavičkou společnosti HaL Computer Systems a O'Reilly & Associates [126].

V současné době je formát spravován a udržován neziskovým konsorciem OASIS⁵⁹ [127], konkrétně Technickým výborem OASIS pro DocBook (*OASIS DocBook Technical Committee*) [128].

Aktuální verze je označena jako verze 5.0. Formát je formálně deklarován v podobě schématu Relax NG⁶⁰. K dispozici je také v podobě schémat DTD⁶¹ a XML Schema [129], [130], [131].

Vzhledem ke skutečnosti, že pro verzi 5.0 formátu DocBook je k dispozici online verze průvodce [132], vychází následující text nejenom z vlastní specifikace formátu, ale také z tohoto průvodce, který nabízí mj. přehledný seznam prvků formátu.

Z hlediska zachycení citačních dat nabízí formát DocBook širokou škálu možností. Pro seznam použitých zdrojů (či bibliografii) lze použít prvek „bibliography“, jednotlivé záznamy mohou být uvedeny ve výrazně strukturované podobě v prvku „biblioentry“ či v méně strukturované podobě v prvku „bibliomixed“. Jak naznačuje J. Kosek [134], užití prvku „bibliomixed“ je vhodné v případě, kdy je potřeba získat přesnou podobu záznamu s interpunkčními znaménky, aniž by uživatel musel měnit styly.⁶² Pro účely citování v textu lze použít prvek „citation“, popř. „xref“ [134]. Zatímco první zmíněný prvek je přímo určen pro citace v textu, druhý je vymezen obecněji jako křížový odkaz na jinou část dokumentu.

Ve formátu DocBook je možné s citačními daty pracovat velmi sofistikovaně – pro naše účely se proto nyní omezíme pouze na charakteristiku prvku „biblioentry“. Tomuto prvku může být podřazeno minimálně dalších 67 prvků.

⁵⁹ OASIS je zkratka názvu *Organization for the Advancement of Structured Information Standards* (Organizace pro rozvoj standardů pro strukturované informace).

⁶⁰ Relax NG znamená *REgular LAnguage for XML Next Generation* (regulární jazyk pro XML příští generace).

⁶¹ DTD je zkratka výrazu *Document Type Definition* (definice typu dokumentu).

⁶² Pro práci s citačními daty ve formátu DocBook lze také použít citační manažery, které tento formát podporují. Patří mezi ně např. manažer RefDB [133], [134].

Z nich 52 uvedeno v následujícím přehledu:

abstract	confgroup	personblurb
address	contractnum	personname
annotation	contractsponsor	printhistory
artpagenums	copyright	productname
author	cover	productnumber
authorgroup	edition	pubdate
authorinitials	editor	publisher
bibliocoverage	extendedlink	publishername
biblioid	issuenum	releaseinfo
bibliomisc	itermset	revhistory
bibliomset	keywordset	seriesvolnums
bibliorelation	legalnotice	subjectset
biblioset	mediaobject	subtitle
bibliosource	org	title
citebiblioid	orgname	titleabbrev
citerefentry	othercredit	volumenum
citetitle	pagenums	
collab	person	

Zbýlých 14 prvků spadá do skupiny prvků označených jako řádkové prvky využitelné při publikování přiřazující význam částem textu v odstavcích (*publishing inlines*) [135]:

abbrev	(db.emphasis)	glossterm
acronym	firstterm	phrase (db.phrase)
coref	footnote	quote
date	footnoteref	subscript
emphasis	foreignphrase	superscript

Prvku „biblioentry“ lze podřadit i další prvky z jiných jmenných prostorů [136].

U prvku „biblioentry“ mohou být využity tzv. společné atributy (*common attributes*) a společné odkazovací atributy (*common linking attributes*).

Způsob práce s prvky a atributy je definován prostřednictvím schématu Relax NG (popř. nenormativních verzí schémat DTD a XML Schema). Jako výhodná se přitom jeví skutečnost, že schéma neurčuje pouze pravidla týkající se použití jednotlivých prvků a atributů, ale u jednotlivých prvků, atributů a hodnot atributů jsou také uváděny základní vysvětlivky (v prvku „a:documentation“⁶³).

Formát DocBook v řadě případů využívá slovníky hodnot atributů. Jako názorný příklad můžeme uvést možné hodnoty atributu „enumeration“ u prvku „biblioid“ [137] (viz obr. 8).

```
<define name="db.biblio.class.enumeration">
  <choice>
    <value>doi</value>
    <a:documentation>A document object identifier.</a:documentation>
    <value>isbn</value>
    <a:documentation>An international standard book number.</a:documentation>
    <value>ismn</value>
    <a:documentation>An international standard technical report number (ISO 10444).</a:documentation>
    <value>issn</value>
    <a:documentation>An international standard serial number.</a:documentation>
    <value>libraryofcongress</value>
    <a:documentation>A Library of Congress reference number.</a:documentation>
    <value>pubsnumber</value>
    <a:documentation>A publication number (an internal number or possibly organizational
      standard).</a:documentation>
    <value>uri</value>
    <a:documentation>A Uniform Resource Identifier</a:documentation>
  </choice>
</define>
```

Obr. 8: Část příkladu schématu Relax NG formátu DocBook týkající se atributu „enumeration“

Je zřejmé, že formát DocBook pracuje s vlastní hierarchií prvků obsahujících citační data. Jako příklad lze uvést autorské údaje – prvku „author“ lze podřadit mj. prvky „personname“ a „affiliation“. Prvek „personname“ pak může být rodičem mj. prvků „honorific“, „firstname“, „surname“ a „othername“. Prvku „affiliation“ mohou být podřazeny mj. prvky „shortaffil“, „jobtitle“, „orgname“ a „orgdiv“ (viz obr. 9).

⁶³ Prefix „a“ ve schématu Relax NG formátu DocBook slouží jako zkrácený výraz pro jmenný prostor <http://relaxng.org/ns/compatibility/annotations/1.0>.

```

<author>
  <personname>
    <honorific>Mr</honorific>
    <firstname>Norman</firstname>
    <surname>Walsh</surname>
    <othername role='mi'>D</othername>
  </personname>
  <affiliation>
    <shortaffil>ATI</shortaffil>
    <jobtitle>Senior Application Analyst</jobtitle>
    <orgname>ArborText, Inc.</orgname>
    <orgdiv>Application Development</orgdiv>
  </affiliation>
</author>

```

Obr. 9: Část průvodce věnovaná prvku „author“

3.3.2 OpenDocument Format for Office Applications

Open Document Format for Office Applications (také zkráceně označovaný jako OpenDocument) je formát, který vznikl jako formát pro reprezentaci elektronických dokumentů. Specifikaci formátu připravila firma Sun Microsystems, v současné době je již tento formát – stejně jako formát DocBook – rozvíjen pod patronací konsorcia OASIS, konkrétně v Technickém výboru OASIS pro Otevřený formát dokumentu pro kancelářské aplikace (OpenDocument) (*OASIS Open Document Format for Office Applications (OpenDocument) Technical Committee*). Práce na formátu byly zahájeny v roce 1999. V roce 2006 byla verze 1.0 formátu schválena ISO a IEC⁶⁴ jako standard ISO/IEC 26300:2006 [138].

V současné době Technický výbor OASIS pro Otevřený formát dokumentu pro kancelářské aplikace (OpenDocument) pracuje na verzi formátu 1.2 [139], jehož některé části jsou dosud v připomínkovém řízení. V následujícím textu je charakterizována tato připravovaná verze formátu (zejména její část 1), neboť předchozí verze formátu 1.1. byla schválena již roku 2007 (pouze konsorciem OASIS, nikoliv již ISO/IEC) a nově připravovaná verze obsahuje řadu podstatných vylepšení.

Ve formátu se vyskytuje řada prvků a atributů, které lze využít ke strukturované práci s citačními daty. Patří mezi ně zejména prvky

⁶⁴ IEC je zkratka názvu *International Electrotechnical Committee* (Mezinárodní elektrotechnická komise).

„text:bibliography“, „text:bibliography-mark“, „text:bibliography-source“, „text:bibliography-entry-template“, „text:bibliography-configuration“ a atributy „text:bibliography-data-field“ a „text:bibliography-type“.

Prvek „text:bibliography-mark“ nemá žádné podřazené prvky, může však být doplněn o celkem 32 atributů:

text:address	text:issn
text:annotate	text:journal
text:author	text:month
text:bibliography-type	text:note
text:booktitle	text:number
text:chapter	text:organizations
text:custom1	text:pages
text:custom2	text:publisher
text:custom3	text:report-type
text:custom4	text:school
text:custom5	text:series
text:edition	text:title
text:editor	text:url
text:howpublished	text:volume
text:identifier	text:year
text:institution	
text:isbn	

Prvek „text:bibliography“ představuje seznam použité literatury vytvořený z jednotlivých prvků „text:bibliography-mark“. Jedním ze dvou prvků, které mu mohou být podřazené, je prvek „text:bibliography-source“, který specifikuje způsob generování seznamu použité literatury, druhým je pak prvek „text:index-body“. Prvku „text:bibliography-source“ mohou rovněž být podřazené dva prvky, z toho se v jednom případě jedná o prvek „text:bibliography-entry-template“ specifikující formátování jednotlivé položky a ve druhém případě o prvek „text:index-title-template“. Prvek „text:bibliography-entry-template“ může mít tři podřazené prvky,

jimiž jsou „text:index-entry-bibliography“, „text:index-entry-span“ a „text:index-entry-tab-stop“.

Spolu s prvkem „text:index-entry-bibliography“ je možné použít atributy „text:bibliography-data-field“ a „text:style-name“. Atributu „text:bibliography-data-field“ má přesně definované možné hodnoty – jedná se o stejnou množinu 32 hodnot, kterou je možné použít jako hodnotu atributu u prvku „text:bibliography-mark“.

Vzhledem ke skutečnosti, že verzi 1.2 formátu OpenDocument již podporují i nové verze kancelářského balíku OpenOffice.org (od verze 3.0 dále [140]), lze určitou představu o uplatnění výše zmíněných atributů získat i přímo v tomto kancelářském balíku. Dokládá to obr. 10, který je snímkem části dialogového okna nástroje Seznam použité literatury).

Obr. 10: Pole pro zachycení položky použité literatury v editoru OpenOffice.org Writer

Pokud se týká atributu „text:bibliography-type“, lze rovněž konstatovat skutečnost, že je možné použít pouze hodnoty spadající do množiny 22 definovaných hodnot:

article	custom4	manual
book	custom5	masterthesis
booklet	email	misc
conference	inbook	phdthesis
custom1	incollection	proceedings
custom2	inproceedings	techreport
custom3	journal	unpublished

www

S výjimkou hodnot „conference“, „custom1“, „custom2“, „custom3“, „custom4“, „custom5“, „email“ a „www“ se jedná o typy dokumentů převzaté z původní specifikace formátu BibTeX (viz dále).

Z výše uvedeného vyplývá, že slovníky hodnot jsou u atributů využívány. Hierarchie prvků s citačními daty není příliš aplikována, neboť je preferováno upřesnění typu dat prostřednictvím atributů u prvku „text:index-entry-bibliography“ bez dalšího členění.

3.3.3 EPUB

Formát EPUB je XML formát určený pro digitální publikace (zejména pro knihy), který zaštiťován organizací IDPF⁶⁵. Byl vytvořen na základě formátu OEB⁶⁶ z roku 1999.

Formát EPUB je složen ze tří částí, jimiž jsou OPS, OPF a OCF⁶⁷. Aktuální verze OPS a OPF jsou verze 2.0 z roku 2007⁶⁸, v případě OCF se jedná o verzi 1.0 z roku 2006 [143].

Pokud se týká zachycení citačních dat, lze konstatovat, že formát EPUB se o nich explicitně nezmiňuje.⁶⁹

V části OPF [142, zejména kap. 2.2] je však věnována pozornost metadatům o celé publikaci. K jejich zachycení je používán prvek „metadata“, který může obsahovat prvky s metadaty tvořícími součást Dublinského jádra, popř. i s dalšími metadaty. Povinně jsou uváděny prvky „title“, „identifier“ a „language“ (s prefixem „dc“) – každý z těchto prvků se musí v daném XML souboru ve formátu EPUB (resp. v části OPF) vyskytovat alespoň jednou. Specifikace části OPF obsahuje podrobný komentář týkající se jednotlivých prvků Dublinského jádra a požadavky či

⁶⁵ IDPF je zkratka názvu *International Digital Publishing Forum* (Mezinárodní fórum pro digitální publikování).

⁶⁶ OEB je zkratka názvu *Open eBook Publication Structure* (otevřená struktura pro publikaci e-knih).

⁶⁷ OPS je zkratka výrazu *Open Publication Structure* (otevřená struktura publikace), OPF znamená *Open Packaging Format* (otevřený balicí formát) a OCF pak *Open Container Format* (otevřený kontejnerový formát).

⁶⁸ V rámci vlastních specifikací těchto formátů je upřesněno, že se jedná o verze 1.0 v rámci verzí 2.0 [141], [142].

⁶⁹ K jejich zahrnutí lze však teoreticky využít tzv. ostrovy XML (*XML Islands*), které představují metodu pro zahrnutí generického obsahu v jazyce XML do formátu EPUB, resp. do jeho části OPS [141].

doporučeními pro jejich implementaci konkrétně v tomto XML formátu. Vzhledem k potřebě zachytit podrobněji roli přispěvatele je pro hodnoty atributu „role“ prvku „contributor“ (opět s prefixem „dc“) určen jako závazný seznam kódů relátorů formátu MARC [144]. Rovněž např. u prvku „format“ je jako nejlepší praktika doporučeno používat hodnoty z řízeného slovníků. Jako příklad jsou zmíněny typy MIME⁷⁰.

3.3.4 TEI

Formát TEI⁷¹ má své počátky již v roce 1987. V roce 1999 pak bylo vytvořeno Konsorcium TEI (*TEI Consortium*). V roce 2002 následovala první verze formátu TEI v jazyce XML – jednalo se o verzi P4, jež v zásadě představovala XML verzi předchozí verze formátu bez větších úprav. Aktuální (od roku 2007) je verze formátu P5 [145].

Formát TEI v této poslední verzi je podrobně popsán v dokumentu *TEI: P5 Guidelines* [146] (pro náš účel je nejpodstatnější kapitola 3.11 tohoto dokumentu). Schémata (konkrétně DTD a Relax NG) jsou k dispozici na platformě SourceForge.net [147].

Pro zachycení citačních dat je možné použít prvky „bibl“, „biblStruct“, „biblFull“ a „listBibl“. Zatímco prvek „bibl“ je určen pro pouze minimálně strukturovaná data (je možné jej také využít pro označení citace v textu), prvek „biblStruct“ obsahuje strukturovaná data (jeho součástí mohou být pouze podřazené prvky bibliografického charakteru, a to ve specifikovaném pořadí). Prvek „biblFull“ je tvořen plně strukturovanými daty, v nichž nejsou použity pouze prvky bibliografického charakteru, ale i další prvky z formátu. Prvek „listBibl“ pak obsahuje seznam záznamů různého druhu (mohou mu být podřazeny např. právě prvky „bibl“ a „biblStruct“). Všechny dosud zmíněné prvky patří do třídy „model.bibLike“.

Ve formátu TEI se dále setkáme s rozlišením bibliografických úrovní citovaných dokumentů (pomocí prvků „analytic“, „monogr“ a „series“). Jedná se o jeden z důkazů silného ovlivnění praxí v oblasti knihovnictví a bibliografie.

⁷⁰ MIME je zkratka výrazu *Multipurpose Internet Mail Extensions* (víceúčelová rozšíření internetové pošty).

⁷¹ TEI je zkratka názvu *Text Encoding Initiative* (Iniciativa pro kódování textu).

Pokud se týká použití atributů, lze konstatovat, že formát TEI s nimi v oblasti citačních dat pracuje minimálně (lze zmínit např. atribut „xml:id“ u prvku „bibl“, který je uveden i v dokumentaci verze P5 formátu). Hierarchické uspořádání dat je využíváno spíše méně, rovněž jednotlivé údaje (např. jméno autora či údaje o rozsahu) nejsou dále rozčleněny. Navíc formát umožňuje značnou flexibilitu, resp. volnost, pokud se týká konzistence uváděných údajů. Tyto skutečnosti názorně dokládá část příkladu z dokumentace, z něhož je zároveň možné získat určitou představu o způsobu označení prvků podřazených výše zmíněným prvkům „biblStruct“ a „bibl“ (viz obr. 11). Další podrobnosti o jednotlivých prvcích jsou k dispozici v rámci dokumentace k verzi P5 TEI [148].

```

<listBibl>
  <head>Bibliography</head>
  <biblStruct xml:id="NELSON80">
    <analytic>
      <author>Nelson, T. H.</author>
      <title>Replacing the printed word:
        a complete literary system.</title>
    </analytic>
    <monogr>
      <title>Information Processing '80: Proceedings of the IFIPS
        Congress, October 1980</title>
      <editor>Simon H. Lavington</editor>
      <imprint>
        <publisher>North-Holland</publisher>
        <pubPlace>Amsterdam</pubPlace>
        <date>1980</date>
      </imprint>
      <biblScope>pp 1013-23 </biblScope>
    </monogr>
    <note>Apparently a draft of section 4 of
      <title>Literary Machines</title>.</note>
  </biblStruct>
  <bibl xml:id="NELSON88">Ted Nelson: <title>Literary Machines</title>
    (privately published, 1987)</bibl>
  <bibl xml:id="BAXTER88">
    <author>Baxter, Glen</author>
    <title>Glen Baxter His Life: the years of struggle</title>
    London: Thames and Hudson, 1988.
  </bibl>

```

Obr. 11: Část příkladu užití prvků „listBibl“, „biblStruct“, „bibl“ a dalších z dokumentace verze P5 formátu TEI

3.3.5 CML

Počátky formátu CML⁷² sahají do 1. poloviny 90. let 20. století, verze 1.0 specifikace CML pak byla publikována v roce 1999 [149]. Přehledná prezentace této verze CML je k dispozici na serveru Zvon.org [150]. V současné době je aktuální verze 2.5 (release candidate).

Aktuality týkající se CML byly publikovány prostřednictvím blogu Petera Murraye-Rusta [151], který stál spolu s Henrym S. Rzepou u zrodu CML; nově je připravován blog přímo zaměřený na CML. Vlastní schéma a další doprovodné informace jsou pak zveřejňovány prostřednictvím platformy SourceForge.net [152], [153].

K zaznamenání citačních dat lze v CML použít atributy prvku „metadata“ a jemu nadřazeného prvku „metadataList“. U prvku „metadata“ je možné uvést metadata založená na Dublinském jádru a metadata specifická pro CML. U prvku „metadata“ je možné použít atributy „content“, „convention“, „dictRef“ a „id“, u prvku „metadataList“ pak prvky „id“, „title“, „convention“, „dictRef“, „name“ a „role“. Ve schématu jsou také specifikovány možné hodnoty atributu „name“ u prvku „metadata“ (jedná se o „metadataType“):

dc:coverage	dc:title	dc:date
dc:description	dc:type	cmlm:safety
dc:identifier	dc:contributor	cmlm:insilico
dc:format	dc:creator	cmlm:structure
dc:relation	dc:publisher	cmlm:reaction
dc:rights	dc:source	cmlm:identifier
dc:subject	dc:language	other

Jedná se tedy o 15 hodnot prvků Dublinského jádra, pět hodnot s prefixem „cmlm“ a o hodnotu nazvanou „other“. U všech zmíněných hodnot je ve schématu doplněn popis, který upřesňuje očekávaný způsob jejich praktického použití.

Názorný příklad, jak mohou vypadat metadata v CML, uvádí obr. 12 (u prvku „metadata“ jsou využity atributy „id“, „name“ a „content“).

⁷² CML je zkratka výrazu *Chemical Markup Language* (chemický značkovací jazyk).

```

<?xml version="1.0" standalone="yes"?>
<metadataList title="metadataList1 example"
xmlns="http://www.xml-cml.org/schema"
xsi:schemaLocation="http://www.xml-cml.org/schema ../schema.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
>
  <metadata id="m1" name="dc:coverage" content="Europe"/>
  <metadata id="m2" name="dc:description" content="Ornithological chemistry"/>
  <metadata id="m3" name="dc:identifier" content="ISBN:1234-5678"/>
  <metadata id="m4" name="dc:format" content="printed"/>
  <metadata id="m5" name="dc:relation" content="abc:def123"/>
  <metadata id="m6" name="dc:rights" content="licence:GPL"/>
  <metadata id="m7" name="dc:subject" content="Informatics"/>
  <metadata id="m8" name="dc:title" content="birds"/>
  <metadata id="m9" name="dc:type" content="bird books on chemistry"/>
  <metadata id="m10" name="dc:contributor" content="Tux Penguin"/>
  <metadata id="m11" name="dc:creator" content="author"/>
  <metadata id="m12" name="dc:publisher" content="Penguinone publishing"/>
  <metadata id="m13" name="dc:source" content="penguinPub"/>
  <metadata id="m14" name="dc:language" content="en-GB"/>
  <metadata id="m15" name="dc:date" content="1752-09-10"/>
  <metadata id="m16" name="cmlm:safety" content="mostly harmless"/>
  <metadata id="m17" name="cmlm:insilico" content="electronically produced"/>
  <metadata id="m18" name="cmlm:structure" content="penguinone"/>
  <metadata id="m19" name="cmlm:reaction" content="synthesis of penguinone"/>
  <metadata id="m20" name="cmlm:identifier" content="smiles:O=C1C=C(C)C(C)C(C)=C1"/>
  <metadata id="m21" name="foo:institution" content="abc.org"/>
  <metadata id="m22" name="foo:bar" content="xyzy"/>
</metadataList>

```

Obr. 12: Příklad example1.xml ze specifikace CML

Jedná se tedy o relativně strukturovaná data, hierarchie je však využita minimálně (jedná se pouze o podřazení prvku „metadata“ prvku „metadataList“).

3.3.6 ThermoML

Formát ThermoML⁷³ byl vytvořen v rámci projektu XML standard IUPAC pro experimentální a kriticky vyhodnocené uložení a zachycení dat o termodynamických vlastnostech (*XML-based IUPAC Standard for Experimental and Critically Evaluated Thermodynamic Property Data Storage and Capture*), který byl realizován v letech 2002-2006 jako projekt IUPAC 2002-055-3-024. Jednalo se o projekt výboru CPEP⁷⁴ [154].

V současné době je aktuální verze 3.0 formátu, která je definována prostřednictvím XML schématu [155]. Podrobné pojednání o formátu nabízí článek [156]. Přehled o dalších zdrojích týkajících se formátu je k dispozici na webu NIST [157]. Podrobnosti o projektu, v jehož rámci byl formát připraven, přináší již

⁷³ ThermoML je zkratka výrazu *Thermodynamics Markup Language* (značkovací jazyk pro termodynamiku).

⁷⁴ CPEP je zkratka názvu *Committee on Printed and Electronic Publications* (Výbor pro tištěné a elektronické publikace).

citovaná stránka z webu IUPAC [154]. Tyto základní zdroje byly využity při hledání odpovědí na stanovené otázky.

Formát ThermoML je složen ze čtyř hlavních bloků, z nichž jedním je blok nazvaný „Citation“ určený k popisu zdroje dat [156, s. 545]. Pracuje jednak s jednoduchými prvky, které nemají žádné podřazené prvky, jednak s prvky komplexními, které jsou doplněny o prvky na nižší hierarchické úrovni.

Do první skupiny prvků patří 18 prvků:

eType	sTitle	sIDNum
eSourceType	sAbstract	sLocation
sDocumentOrigin	sKeyword	sVol
sAuthor	eLanguage	sPage
sPubName	sDOI	
yrPubYr	urlCit	
dateCit	sCASCit	

U prvků, jejichž název začíná písmenem „e“, lze ve formátu ThermoML používat pouze hodnoty z množiny hodnot uvedených ve schématu (písmeno „e“ odkazuje na výraz „enumeration“, tj. výčet). Z uvedených 18 prvků jsou tři výčtové („eType“, „eSourceType“ a „eLanguage“).

Možných hodnot prvku „eType“ je pouze deset:

book	conferenceProceedings
journal	archivedDocument
report	personalCorrespondence
patent	publishedTranslation
thesis	unspecified

Prvek „eSourceType“, který je na rozdíl od prvku „eType“ zaměřen na status citovaného zdroje, může dosáhnout pouze tří hodnot, a to „Original“, „ChemicalAbstracts“ nebo „Other“.

U prvku „eLanguage“ se jedná o osm hodnot:

Chinese	German
English	Japanese
French	Polish

Russian

Other Language

Prvky s názvem začínajícím na písmeno „s“ jsou prvky, jejichž hodnotou je řetězec (*string*). Jedná se o celkem dvanáct prvků z výše uvedeného přehledu. Prvky, jejichž označení začíná výrazem „date“, jsou prvky s daty – z množiny 18 prvků se jedná pouze o jeden prvek. Stejně tak pouze jeden prvek má název začínající výrazem „url“, který naznačuje očekávanou přítomnost webové adresy.

Komplexními prvky jsou pak čtyři prvky:

TRCRefID	journal
book	thesis

Prvku „TRCRefID“⁷⁵ jsou podřazeny prvky „yrYRPub“, „sAuthor1“, „sAuthor2“ a „nAuthorn“, prvku „book“ pak prvky „sChapter“, „sEdition“, „sEditor“, „sISBN“ a „sPublisher“, prvku „journal“ prvky „sISSN“ a „sCODEN“⁷⁶ a prvku „thesis“ prvky „sDeg“, „sDegInst“ a „sUMIPubNum“.

Vysvětlivky k jednotlivým prvkům jsou obsaženy jako hodnota prvku „xsd:documentation“ ve schématu.

Přestože formát ThermoML obsahuje řadu vlastních prvků využitelných pro účely citování, ve specifikaci nejsou dostatečně uvedeny instrukce, jak s prvky pracovat. Např. u prvku „sAuthor“ není zřejmé, v jaké podobě má být jména autora uvedeno, tj. zda je upřednostňována invertovaná či přímá podoba, zda je možné uvádět úplné znění či pouze iniciály prvního jména atd. Citační data tedy v tomto formátu nejsou vždy dostatečně strukturována.

3.3.7 Interní XML formát Zlaté knihy

Interní XML formát využitý ve Zlaté knize (resp. část tohoto formátu zahrnující citační data) je popsán na základě části zdrojových dat, na jejichž základě byla vytvořena pokročilá XML verze Zlaté knihy, která je dostupná na webu IUPAC [159]. Tato data byla získána od B. Košaty [9].

Formát byl vyvinut v rámci projektu Standardní XML datové slovníky pro chemii (*Standard XML data dictionaries for chemistry*) [160].

⁷⁵ Jedná se o identifikátor referenčních dat TRC (*Thermodynamics Research Center*, Výzkumné centrum pro termodynamiku [158]).

⁷⁶ CODEN je zkratka výrazu *Code Number*, tj. kódové číslo. CODEN slouží k identifikaci názvů seriálů, popř. i jiných typů publikací.

Všechny záznamy jsou podřazeny prvku „sources“ – údaje o jednotlivých zdrojích jsou k dispozici v prvcích označených jako „source“. Zdroje jsou pak rozčleněny do několika skupin odpovídajících skupinám záznamů v online verzi Zlaté knihy, která je na základě těchto zdrojových dat vygenerována.

Citační data ze skupiny označené v online verzi Zlaté knihy výrazem „Books“ jsou v XML formátu uložena v prvku „source“ bez jakýchkoliv atributů. Prvku „source“ jsou v tomto případě podřazeny prvky „title“, „nickname“ a „abbrev“. Jako hodnota prvku „title“ je vždy uveden nejenom název knihy, ale také nakladatelské údaje o knize, popř. upřesnění částí knihy (příklad viz obr. 13).

```
<source>
  <title>IUPAC Nomenclature of Organic Chemistry. Sections A, B, C, D, E, F and H,
    Pergamon Press, Oxford, 1979.</title>
  <nickname>Blue Book</nickname>
  <abbrev>B.B.</abbrev>
</source>
```

Obr. 13: Příklad údajů o knize v interním XML formátu Zlaté knihy

Jako hodnota prvku „nickname“ je uváděno zkrácené označení daného zdroje, které je v online verzi využito jako vstupní prvek záznamu. Hodnota prvku „abbrev“ pak byla využita jako identifikátor citací zdrojů u jednotlivých hesel ve Zlaté knize (doplněná o konkrétní citovanou stránku).

V případě skupiny zdrojů nazvané „ISO“ je u prvku „source“ vždy uveden atribut „type“ s hodnotou „ISO“. Podřazenými prvky prvku „source“ v tomto případě jsou prvky „y“, „number“ a „title“ udávající rok vydání dané normy, její číselné označení a celý název (včetně označení a názvu příslušné části). Jako identifikátor citací zdrojů u jednotlivých hesel jsou využity hodnoty prvků „number“ a „y“, jimž je nadřazen prvek „source“ s atributem „id“ s hodnotou „ISO“.

U skupiny zdrojů „Pure and Applied Chemistry“ je u prvku „source“ opět využit atribut „type“, tentokrát s hodnotou „PAC“. Prvku „source“ jsou u této skupiny zdrojů podřazeny prvky „y“, „v“, „p“, „endp“, „title“ a „link“, jejichž hodnotami jsou rok vydání daného článku, ročník časopisu, číslo stránky, na němž článek začíná, číslo poslední stránky článku, název článku a prázdný prvek. U prvku „link“ jsou totiž využity atributy „href“, jehož hodnotou je URL, které odkazuje na úplný text článku ve formátu PDF na webu IUPAC, a „doi“, které má vždy hodnotu

identifikátoru DOI přiřazeného danému článku. K identifikaci citací u hesel jsou využity hodnoty prvků „y“ a „v“ doplněné o prvek „p“, jehož hodnotou je konkrétní citovaná stránka z daného článku (nadřazeným prvkem je opět prvek „source“).

Poslední skupina zdrojů, která je v online verzi Zlaté knihy označena jako „Other“, obsahuje pouze dva zdroje. První z nich má jako hodnotu atributu „type“ u prvku „source“ uvedenu hodnotu „VIM“ a prvek „source“ má podřazený prvky „title“, „nickname“ a „abbrev“, tj. stejnou sadu prvků, s níž se setkáváme u skupiny „Books“. Druhý záznam má jako hodnotu atributu „type“ u prvku „source“ uvedenu hodnotu „CODATA“, přičemž prvek „source“ je podřazen pouze jeden prvek, a to prvek „title“. Obsahuje bibliografické údaje o časopiseckém článku, tj. nikoliv pouze název, jak by napovídalo označení prvku. Navíc jsou v rámci tohoto prvku využity i prvky „i“ sloužící v online verzi Zlaté knihy k vizuálnímu odlišení názvu a ročníku časopisu od zbývajících údajů formou kurzivy. V prvním případě je jako identifikátor citací zdroje použit výraz „CODATA Bull.“ uvedený jako hodnota atributu „id“ u prvku „source“ (dále jsou doplněny hodnoty prvků „y“, „v“ a „p“). Ve druhém případě je jako identifikátor použit pouze prázdný prvek „source“ s hodnotou „VIM“ atributu „id“.

3.4 Dílčí závěry

S rozvojem sémantického webu se objevuje řada projektů zaměřených na zlepšení reprezentace citačních dat ve strojově srozumitelné podobě. Patří mezi ně jazyk CSL sloužící k formátování citací a záznamů v různých citačních stylech, ontologie BIBO a CiTO věnující se specifikaci pojmů spojených s citováním a jejich vztahů či projekt OObib, jehož cílem je zlepšit citační funkce kancelářského balíku OpenOffice.org.

Paralelně je vytvářena řada XML formátů, jejichž primárním cílem není zachycení citačních dat. Mezi ně patří formáty DocBook, OpenDocument Format for Office Applications, EPUB, TEI, CML, ThermoML či interní XML formát Zlaté knihy. Některé z těchto formátů mají velmi podrobně zpracovanou oblast citačních dat – patří mezi ně například formáty DocBook či TEI. Naopak formát EPUB se citačními daty příliš nezabývá.

4. Citační formáty

4.1 Úvod

V této kapitole je zachycen přehled vybraných citačních formátů, který obsahuje jak charakteristiku formátu vycházející z jeho specifikace (je-li k dispozici), tak zjištění týkající se implementace formátu v praxi v různých elektronických informačních zdrojích.

Citační formáty byly vybrány na základě průzkumu své užívanosti v praxi [5], a to v elektronických informačních zdrojích, které jsou uvedeny následující podkapitole. Do přehledu byly zařazeny pouze formáty prezentující data ve strukturované podobě, a to jak formáty nevyužívající jazyk XML, tak formáty na tomto jazyce založené. Podmínkou zařazení formátu se stala možnost rozlišit v záznamu jednotlivé datové prvky (údaje), resp. jejich dvě části, tj. identifikaci (jméno) údaje a hodnotu údaje (viz [161, s. 88-89]). V případě formátů nevyužívajících jazyk XML jsou s ohledem na srozumitelnost a čtivost v dalším textu kromě výše uvedených výrazů používány také výrazy název pole a obsah pole (nebo synonymně hodnota pole).

Do první skupiny byly zařazeny formáty RIS, ISI Export Format, RefWorks, EndNote, BibTeX a OpenURL COinS⁷⁷, do druhé pak formáty MARCXML⁷⁸, MODS⁷⁹ a NLM⁸⁰.

Naopak záměrně nebyly zahrnuty formáty s nestrukturovanými daty, které některé zdroje nabízejí. Jedná se o souvislý text bez explicitního označení typů jednotlivých položek – svou podobou proto tyto formáty spíše připomínají citační styly.

Příklady formátů s nestrukturovanými daty jsou pro názornost uvedeny na obr. 14 a 15. Obr. 14 byl pořízen v digitální knihovně ScienceDirect – je zvolen

⁷⁷ COinS je zkratka výrazu *ContextObject in SPAN* (kontextové objekty mezi značkami „SPAN“).

⁷⁸ MARCXML je zkratka výrazu *MARC Extensible Markup Language* (značkovací jazyk pro formát MARC).

⁷⁹ MODS je zkratka výrazu *Metadata Object Description Schema* (schéma pro popis objektů metadaty).

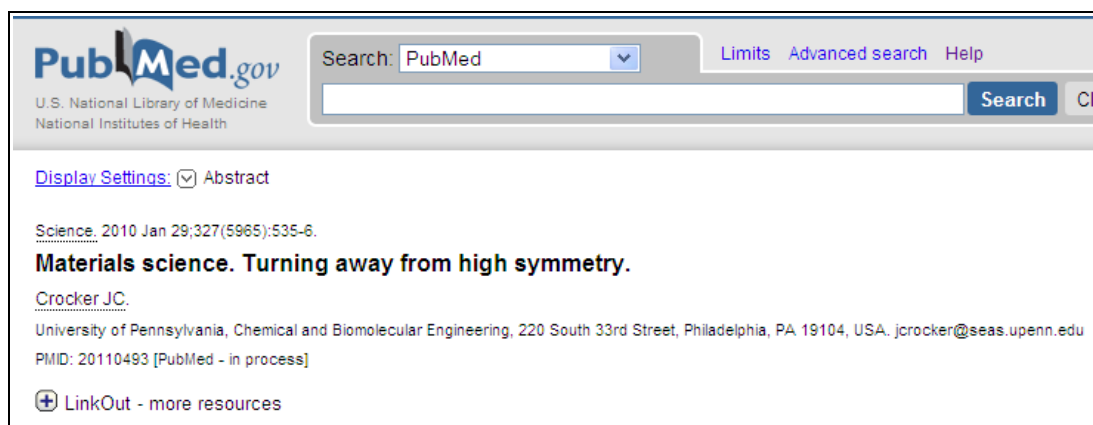
⁸⁰ NLM je zkratka názvu *National Library of Medicine* (Národní lékařská knihovna).

náhled prostřednictvím internetového prohlížeče, neboť zdrojový kód dané HTML stránky byl velmi komplikovaný a nepřehledný.

```
Mary S. Woodley, A digital library project on a shoestring, Library Collections, Acquisitions,
(http://www.sciencedirect.com/science/article/B6VSH-46WMTF3-3/2/91743f65924d52174b81a076c9f501
Abstract:
Supported by a modest grant and collaboration with local historical societies, the California
Keywords: Digital library; Digital archive; Metadata standards and workflow
```

Obr. 14: Část záznamu časopiseckého článku z digitální knihovny ScienceDirect ve formátu ASCII

Obr. 15 byl převzat ze systému PubMed a zobrazen přímo v nativním grafickém rozhraní systému.



Obr. 15: Záznam časopiseckého článku ze systému PubMed ve formátu PubMed Citation

Podrobnější přehled všech citačních formátů, do nichž je možné exportovat data ze zvolených informačních systémů, je podán v přílohách 1 až 14.⁸¹

Nyní se vraťme k citačním formátům se strukturovanými daty. U nich je pozornost zaměřena na čtyři klíčové oblasti, jimiž jsou specifikace, syntax, sémantika a pravidla pro tvorbu obsahu. Na závěr je v případě potřeby zařazena část věnovaná poznámkám, v níž je možné zachytit případné další doplňující údaje o formátu (např. u formátů nevyužívajících jazyk XML o existenci verze formátu v tomto jazyce).

Aby bylo možné přehled citačních formátů efektivně využít v další práci, byly u jednotlivých klíčových oblastí stanoveny rámcové otázky, na něž je u jednotlivých formátů zjišťována konkrétní odpověď. Jednalo se o následující otázky:

⁸¹ Doplňkově jsou zařazeny příklady exportů přímo do citačních stylů (pouze v případech, kdy je daný informační systém nabízí).

Specifikace

- Je k dispozici specifikace formátu? Pokud ano, kde a v jaké podobě?
- Jsou zveřejňovány údaje o aktualizaci formátu? Pokud ano, kde a v jaké podobě?

Syntax

- Je jednoznačně určeno kódování?
- Jak je rozlišen začátek a konec záznamu v daném formátu?
- Jak formát pracuje s tzv. bílými znaky?
- Jaký má ve formátu význam konec řádku?
- Jak formát řeší řádkování (např. v abstraktech)?
- Co se stane v případě více hodnot jednoho (typu) údaje?

Sémantika

- Jak jsou označovány jednotlivé typy údajů?
 - o Jsou rozlišováni editoři a autoři?
 - o Jsou rozlišovány zkrácené názvy časopisů od názvů úplných?
- Kolik prvků formát obsahuje?
- O které konkrétní prvky se jedná?

Pravidla pro tvorbu obsahu

- Jsou předepsána pravidla pro vyplňování údajů (obsahu polí)? Pokud ano, jsou jednoznačná, nebo umožňují variabilitu?

Poznámky

- U formátů nezaložených na jazyce XML:
 - o Existuje oficiální verze formátu v jazyce XML? Pokud ano, je veřejně dostupné její specifikace?
- Podle potřeby další doplňující údaje⁸²

Pokud se týká zdrojů použitých při zpracování přehledu, jsou jimi dokumenty uvedené v části Specifikace.

⁸² Tato položka v části věnované poznámkám nemá podobu otázky, je zde však pro úplnost uvedena.

Seskupení rámcových otázek do bloků nazvaných Syntax, Sémantika a Pravidla pro tvorbu obsahu je inspirováno členěním, které nabízí Caplanová [162, s. 6-7] ve své práci věnované metadatovým schémátům. Tato autorka specifikuje tři aspekty těchto schémat, jimiž jsou právě sémantika (*semantics*), pravidla pro tvorbu obsahu (*content rules*) a syntax (*syntax*). Sémantikou je míněno označení významu položek metadat (metadatových prvků). Pravidla pro tvorbu obsahu udávají, jak jsou hodnoty metadatových prvků vybírány a reprezentovány. Syntax schématu je způsob, jakým mají být jeho prvky zakódovány ve strojově čitelné podobě.⁸³ V naznačeném smyslu je členění využito i v této práci.

Jednotlivé bloky obsahují nejprve údaje zjištěné studiem specifikace formátu, popř. jiných či dalších vhodných zdrojů. Následuje stručný komentář týkající se implementace formátu ve vybraných existujících informačních systémech, a to na základě speciálně vytvořených datových sad (kap. 4.2). V tomto směru hrají též velmi podstatnou roli tzv. bílé znaky, resp. práce s nimi v praxi. S ohledem na skutečnost, že se jedná o jednu z klíčových součástí přehledu, vlastnímu přehledu formátů navíc předchází teoretický úvod týkající se těchto znaků (kap. 4.3).

4.2 Tvorba datových sad

Za účelem získání možnosti ověřit implementace citačních formátů v praxi byly ve vybraných elektronických informačních zdrojích⁸⁴ (dále jsou označovány obecně jako systémy, popř. specifickěji podle svého typu) vytvořeny sady s daty reprezentujícími citační údaje týkající se jednotlivých dokumentů (často časopiseckých článků, kapitol z knih či jiných typů příspěvků) dostupných prostřednictvím těchto zdrojů.

Při výběru systémů bylo postupováno tak, aby se jednalo o vyvážený vzorek reprezentující různé typy informačních zdrojů zejména z oblasti STM, a to z hlediska typů zpřístupňovaných dat či dokumentů, formy zpřístupnění a významu pro vědeckou práci.

⁸³ Z pohledu tématu této práce není bez zajímavosti, že autorka také upozorňuje, že sémantika, syntax a pravidla pro tvorbu obsahu jsou často neúplná, neobsahují mnoho zákazů nebo nabízejí při implementaci řadu možností [162, s. 7].

⁸⁴ Jednalo se o zdroje, jejichž obsah je částečně nebo zcela dostupný zdarma prostřednictvím internetu nebo které jsou (aktuálně) dostupné na Univerzitě Karlově v Praze prostřednictvím Portálu elektronických zdrojů.

Celkem bylo vybráno čtrnáct zdrojů, konkrétně:

- dva nejdůležitější světové systémy zpřístupňující citační data (Web of Science a Scopus),
- dva významné vědecké časopisy s vysokým impakt faktorem (Nature a Science),
- tři komerční digitální knihovny vydavatelů (SpringerLink, ScienceDirect a Wiley InterScience),
- systém zaměřený na oblast věd o životě a biomedicínu (PubMed),
- dva systémy pro zpřístupňování databázových balíků (EBSCOhost a ProQuest Central),
- dva souborné knihovní katalogy (WorldCat a dánský souborný katalog bibliotek.dk),
- dva systémy zaměřené na zpřístupňování e-printů a jiných typů dokumentů, přičemž oba pro svůj provoz využívají software s otevřeným zdrojovým kódem (CERN⁸⁵ Document Server založený na softwaru CDS⁸⁶ Invenio a E-LIS⁸⁷ založený na softwaru EPrints).⁸⁸

Data z těchto zdrojů byla získána v březnu až dubnu 2010, reflektují tedy stav k uvedenému časovému období.

Pro přípravu dat byl (v případech, které to umožňovaly) použit tento postup:

1. V základním (resp. implicitně nastaveném) rozhraní rešeršního systému daného zdroje byl zadán jednoduchý dotaz, a to výraz digital libraries. (Cílem tohoto kroku bylo pouze získat sadu záznamů bez zřetele k tematickému profilu dokumentů, které reprezentují.)
2. Ze zobrazených výsledků bylo vybráno prvních 25 (vzhledem ke smyslu tohoto kroku bylo vždy ponecháno implicitní nastavení zobrazení pořadí výsledků).

⁸⁵ CERN je zkratka pocházející z francouzského výrazu *Conseil Européen pour la recherche nucléaire* (Evropská organizace pro jaderný výzkum).

⁸⁶ CDS je zkrácené vyjádření výrazu *CERN Document Server* (Dokumentový server CERN).

⁸⁷ E-LIS je zkrácený název repozitáře *E-prints in Library and Information Science* (E-printy v knihovní a informační vědě).

⁸⁸ Předpokladem efektivní práce se zdroji dostupnými zdarma je často registrace (rovněž zdarma) – z vybraných zdrojů se jedná o časopisy Nature a Science (resp. jejich webové prezentace) a katalogy WorldCat a bibliotek.dk.

3. Sada 25 záznamů byla vyexportována v citačním formátu, resp. citačních formátech, které byly dostupné a zároveň patřily mezi formáty vybrané k podrobnější charakteristice v této práci.

Ne vždy bylo možné postupovat přesně podle této metodiky. Některé zdroje neumožňovaly export více než jednoho záznamu v jednom souboru (tj. najednou), konkrétně se jednalo o časopisy Nature a Science, o citační databázi Web of Science (pouze v případě formátu BibTeX) a dále o CERN Document Server. V těchto případech byly vytvořeny soubory obsahující vždy pouze první záznam v pořadí.

Naopak e-printový archiv E-LIS umožňoval buď stažení jednotlivých záznamů v citačních formátech, nebo celé množiny výsledků vyhledávání. Nebylo tedy možné pracovat pouze s podmnožinou výsledků. V tomto případě byl z důvodu příliš vysokého počtu získaných výsledků při zadání dotazu digital libraries zvolen dotaz digital libraries software – vytvořená sada výsledků je tvořena celkem 71 položkou. Pokud se týká exportu sady ve formátu RIS, bylo v archivu E-LIS nutné si záznamy v tomto formátu nejprve zobrazit formou HTML stránky a teprve následně uložit jako soubor (již ve formátu RIS). Na HTML stránce došlo k nesprávnému zobrazení některých znaků (např. písmene „á“), přičemž prohlížeč Mozilla Firefox (verze 3.5.8) rozpoznal znakovou sadu jako UTF-8⁸⁹. V některých případech problém s chybným zobrazením přetrvával i po uložení dat ve formátu RIS (např. „Politechnika ?ódzka“). Vzhledem ke skutečnosti, že v archivu E-LIS se liší nabídka formátů, do nichž lze exportovat všechny výsledky vyhledávání, a nabídka formátů, do nichž lze exportovat jednotlivé záznamy, bylo v případě formátů OpenURL COinS a MODS možné vytvořit pouze soubory s jedním záznamem.

Podrobnější komentář si zaslouží i výsledky získané z dokumentového serveru CERN. Již výše bylo konstatováno, že bylo možné získat pouze exporty jednotlivých záznamů. Je však ještě nezbytné doplnit, že záznamy, které nebyly ve formátu v jazyce XML, bylo potřeba uložit manuálně na základě výběru dat z HTML stránky. Za účelem zajištění co nejpřesnějšího převzetí těchto dat byla data nejprve zobrazena ve zdrojovém kódu stránky (tj. ve formátu HTML) a teprve následně

⁸⁹ UTF je osmibitový transformační formát UCS/Unicode (*8-bit UCS/Unicode Transformation Format*).

uložena do souboru (ve formátu TXT⁹⁰ a v kódování UTF-8). U formátu označeného výrazem EndNote se nejednalo o klasickou verzi tohoto formátu, ale jeho verzi v jazyce XML.

V rozhraní systému PubMed bylo implicitně navrženo uložení dat v podobě textového souboru (ve formátu TXT); protože se však jednalo o data v jazyce XML (ve formátu NLM), byla pro další zpracování přípona souboru z pragmatických důvodů změněna z TXT na XML.

V souborných knihovních katalozích WorldCat a bibliotek.dk bylo možné získat sadu dat ve formátu RIS (v případě katalogu WorldCat se jednalo o soubor ve formátu RIS, v případě katalogu bibliotek.dk pak o soubor ve formátu TXT); nabízen byl též formát vhodný pro citační software EndNote, avšak nejednalo se o formát EndNote, ale rovněž o formát RIS, proto není dále uvažován. V prvním případě bylo možné data získat stažením dvou binárně totožných souborů⁹¹, v druhém případě bylo možné si stáhnout pouze jediný soubor.

V digitálních knihovnách SpringerLink a ScienceDirect a v systémech Scopus, EBSCOhost a ProQuest Central bylo možné postupovat přesně v souladu se stanovenou metodikou, nevyskytly se žádné odchylky, které by realizaci naplánovaných kroků významně ztížily či zcela znemožnily.

Podobně tomu bylo v případě digitální knihovny Wiley InterScience – zde je však možné doplnit, že export ve formátu EndNote byl realizován ve třech nabízených variantách pro různé operační systémy, a to pro Microsoft Windows, pro Mac OS a pro unixové operační systémy.

V případě systému ProQuest Central doplňme, že byly vytvořeny dvě sady záznamů ve formátu RIS, přesněji ve formátu RIS a formátu jemu velmi blízkém. Uživatel uvedeného systému má totiž možnost stáhnout si záznamy ve formátu vhodném pro citační softwary ProCite, EndNote a Reference Manager (ke stažení je nabízen soubor ve formátu RIS) a dále ve formátu kompatibilním s citačními softwary ProCite, EndNote, Reference Manager a Refworks (v tomto případě se jedná o formát velmi blízký formátu RIS; uživatel je vyzván k uložení dat v podobě textového souboru).

⁹⁰ TXT je přípona textového souboru (*textfile*).

⁹¹ Shodnost obou souborů byla prověřena v softwaru KDiff3.

U systému EBSCOhost lze ještě doplnit, že přestože nenabízí možnost exportu výsledků ve formátu MARCXML, je možné záznamy získat ve formátu MARC 21. Obdobně systémy CERN Document Server, Science a PubMed umožňují export záznamů (jak již bylo uvedeno, v případě serveru CERN vždy pouze jednotlivých) do formátu Medlars/Medline, tj. vývojového předchůdce formátu NLM. Ty však (stejně jako formát MARC 21 a některé další formáty) nebyly v rámci předkládané práce vybrány k podrobnější analýze.

Obecně lze ještě upřesnit, že kdykoliv byla uvedena možnost výběru míry úplnosti záznamu či záznamů (např. záznam obsahující bibliografické údaje i abstrakt), byla vybrána varianta s neúplněnějšími údaji (jednalo se mj. o systémy Web of Science a Scopus). V případě formátů RefWorks a EndNote (resp. EndNote Web) pak byla stažena data pouze v případě, že bylo možné je uložit formou samostatného souboru, tj. nejednalo se o data určená pro přímý export do citačních manažerů RefWorks a EndNote Web. V tab. 4 jsou případy, kdy byl k dispozici pouze přímý export, označeny kurzivou.

Tab. 4: Přehled formátů a zdrojů dat

Zdroje	Formáty								
	RIS	ISI	RefWorks	End Note	BibTeX	OpenURL COinS	MARC XML	MODS	NLM
Web of Science	–	ano	–	ano	ano	–	–	–	–
Scopus	ano	–	ano	–	ano	–	–	–	–
Nature	ano	–	–	–	–	–	–	–	–
Science	ano	–	ano	ano	ano	–	–	–	–
SpringerLink	ano	–	–	–	–	–	–	–	–
ScienceDirect	ano	–	ano	–	ano	–	–	–	–
Wiley InterScience	–	–	–	ano	–	–	–	–	–
PubMed	–	–	–	–	–	–	–	–	ano
EBSCOhost	ano	–	ano	ano	ano	–	–	–	–
ProQuest Central	ano	–	ano	–	–	–	–	–	–
WorldCat	ano	–	ano	–	–	–	–	–	–
bibliotek.dk	ano	–	ano	–	–	–	–	–	–
CERN Document Server	–	–	–	ano	ano	–	–	ano	ano
E-LIS	ano	–	–	ano	ano	ano	ano	–	–

Na základě dat uvedených v tab. 4 lze shrnout, že v deseti případech ze čtrnácti je nabízen citační formát RIS. Exportní formát ISI (ISI Export Format, dále

jen formát ISI) je k dispozici pouze v systému Web of Science. Citační formát RefWorks je dostupný v sedmi případech, z toho je v šesti případech k dispozici pouze přímý export do softwaru RefWorks. Ten je velmi užitečný pro uživatele, ale (jak již bylo poznamenáno) pro naše účely nikoliv, proto není uvedených šest případů podrobněji rozpracováno. V šesti ze zkoumaných systémů je dostupný formát EndNote, z toho ve dvou případech výhradně v podobě přímého exportu do citačního softwaru EndNote Web (tyto dva případy opět nejsou dále uvažovány). Formát BibTeX je k dispozici v sedmi případech, formát NLM ve dvou případech zbylé formáty, tj. OpenURL COinS, MARCXML a MODS, vždy pouze v jednom případě.

Je tedy zřejmé, že do určité míry již tradiční formáty typu RIS, EndNote či BibTeX jsou ve vybraném vzorku systémů přítomné často, z novějších formátů se jedná především o formát RefWorks. Vybrané formáty v jazyce XML, tj. MARCXML, MODS a NLM, jsou zatím implementovány spíše výjimečně.

Celkem bylo vytvořeno 31 datových sad, přičemž 17 z nich obsahovalo 25 záznamů, ve třech sadách bylo zařazeno po 71 záznamu a v jedenácti sadách po jednom záznamu. Jednotlivým datovým sadám byly přiřazeny identifikátory složené vždy z tříznakové (konkrétně třípísmenné) zkratky citačního formátu, spojovníku a číselného označení (od 1 dále).

Přehled základních údajů o vytvořených datových sadách (včetně jejich identifikátoru) je k dispozici v tab. 5.

Tab. 5: Přehled datových sad

Citační formát	Identifikátor datové sady	Zdroj dat	Počet záznamů	Formát soubor (přípona)	Upřesnění
BibTeX	bib-1	Web of Science	1	bib	–
	bib-2	Scopus	25	bib	–
	bib-3	Science	1	txt	–
	bib-4	ScienceDirect	25	bib	–
	bib-5	EBSCOhost	25	htm	–
	bib-6	CERN Document Server	1	txt	–
	bib-7	E-LIS	71	bib	–
EndNote	end-1	Science	1	enw	–
	end-2	Wiley InterScience	25	enw	pro Microsoft Windows
	end-3	Wiley	25	enw	pro Macintosh

Citační formát	Identifikátor datové sady	Zdroj dat	Počet záznamů	Formát soubor (přípona)	Upřesnění
		InterScience			
	end-4	Wiley InterScience	25	enw	pro Unix nebo Linux
	end-5	CERN Document Server	1	xml	–
	end-6	E-LIS	71	enw	–
ISI	isi-1	Web of Science	25	cgi ⁹²	–
MARCXML	mar-1	CERN Document Server	1	xml	–
MODS	mod-1	E-LIS	1	xml	–
	nlm-1	PubMed	25	xml	–
NLM	nlm-2	CERN Document Server	1	xml	–
OpenURL COinS	ope-1	E-LIS	1	txt	–
RefWorks	ref-1	Science	1	htm	–
	ris-1	Scopus	25	ris	–
	ris-2	Nature	1	ris	–
	ris-3	Science	1	htm	pro Zotero
	ris-4	SpringerLink	25	ris	–
	ris-5	ScienceDirect	25	ris	–
	ris-6	EBSCOhost	25	ris	–
	ris-7	ProQuest Central	25	ris	–
RIS	ris-8	ProQuest Central	25	txt	formát blízký formátu RIS a vhodný pro ProCite, EndNote a Reference Manager
	ris-9	WorldCat	25	ris	–
	ris-10	WorldCat	25	ris	pro EndNote
	ris-11	Bibliotek	25	txt	–
	ris-12	E-LIS	71	ris	–

4.3 Bílé znaky

Jak již bylo naznačeno, podstatnou roli – zejména s ohledem na strojové zpracování – hrají tzv. bílé znaky (*white spaces*). Pro potřeby této práce jsou za bílé

⁹² V databázích dostupných prostřednictvím platformy Web of Knowledge je možné získat záznamy (ať již jednotlivé či jejich množiny) ve formátu ISI prostřednictvím tlačítka „Save to EndNote, RefMan, ProCite“.

znaky pokládány čtyři znaky, které jsou uvedeny v tab. 6⁹³ [163, s. 501], [164, s. 2], [1], [2].

Tab. 6: Přehled vybraných bílých znaků

Reprezentace znaku v kódování Unicode	Hexadecimální reprezentace znaku	Slovní reprezentace znaku v angličtině	Slovní reprezentace znaku v češtině	Zkratka pro znak použitá v této práci
U+000A	0A	line feed	nový řádek	LF
U+000D	0D	carriage return	posun na začátek řádku	CR
U+0009	09	character tabulation	tabulátor	HT
U+0020	20	space	mezera	S

Ve specifikaci formátu XML jsou jako bílé znaky v kap. 2.10 explicitně uvedeny znaky pro mezery, pro tabulátory a prázdné řádky, dále jsou v kap. 2.11 zmíněny též znaky pro nový řádek a posun na začátek řádku [1], [2].

Tzv. nesignifikantní bílé znaky mohou být používány oddělení značek za účelem zvýšení čitelnosti dokumentu – XML procesor je vždy musí vynechat ze zpracování. Naopak tzv. signifikantní bílé znaky musí být uchovány. Způsob zacházení s bílými znaky je možné určit pomocí atributu `xml:space` a jeho hodnot „preserve“ nebo „default“. Jsou-li přítomné konce řádků (typicky naznačené přítomností znaků CR, tj. #xD, a LF, tj. #xA), XML procesor se musí chovat tak, jako kdyby byly všechny tyto znaky normalizovány do podoby jednoho znaku LF (#xA) [1], [2].

Význam a výskyt bílých znaků byl zkoumán jak po stránce teoretické (údaje ve specifikaci citačních formátů), tak po stránce praktické (užití bílých znaků v jednotlivých implementacích formátů). Je nutné zdůraznit, že bílé znaky jsou zkoumány ve vztahu k syntaxi, není tedy např. zkoumán výskyt znaků pro mezery v rámci hodnot polí.⁹⁴

⁹³ Bílé znaky patří mezi tzv. netisknutelné znaky. Netisknutelným znakem je míněn „[n]ealfanumerický znak zobrazující informace týkající se formátování textu (např. tabulátor, konec odstavce, pevná mezera apod.), který není součástí vlastního textu (je viditelný pouze na monitoru)“ [41, heslo netisknutelný znak]. V angličtině je netisknutelný znak označován jako *non-print character*, popř. *non-printable character* (viz např. [165]).

⁹⁴ V dalším textu a v tabulkách jsou v případech, kdy je rozebírána syntax, na pozicích zkoumaných bílých znaků uvedeny barevně zvýrazněné zkratky.

4.4 Formáty nevyužívající jazyk XML

4.4.1 RIS

Specifikace formátu

Specifikace citačního formátu RIS je k dispozici v rámci webové prezentace citačního softwaru Reference Manager firmy Thomson Reuters [166]. Úvodní stránka specifikace slouží jako rozcestník k patnácti dalším stránkám obsahujícím konkrétní údaje o jednotlivých typech polí, jejich názvech (jedná se o tzv. tagy), dále typech dokumentů a rovněž celkem šest příkladů záznamů v tomto formátu.

Pokud se týká aktualizace formátu RIS, lze konstatovat, že na všech šestnácti stránkách dokumentace (úvodní stránka a patnáct odkazovaných stránek) je uvedeno datum jejich poslední úpravy (jedná se o data z let 2001 a 2004).

Syntax

Kromě identifikátoru záznamu, který může pouze obsahovat číslice 0-9 a písmena A-Z, je povoleno užití znaků z rozšířené znakové sady IBM⁹⁵ (*IBM Extended Character Set*), konkrétně od znaku 32 (*space*, mezera) do znaku 255. Určitou výjimku tvoří znak 42 (*asterisk*, hvězdička) – není povoleno užití tohoto znaku v polích s autorskými údaji, klíčovými slovy a názvy periodika.

Praxe však ukazuje, že je v řadě případů potřeba zachytit i znaky, které nejsou součástí rozšířené znakové sady IBM. Dokladem může být skutečnost, že u pěti sad záznamů z dvanácti testovaných textových editorů PSPad identifikoval znakovou sadu jako UTF-8, tj. znaky z rozšířené znakové sady IBM nepostačovaly ke správnému zobrazení obsahu souboru (viz tab. 7).

Tab. 7: Kódování souborů ve formátu RIS

ID ⁹⁶ sady	Kódování rozpoznané editorem PSPad
RIS-1	UTF-8
RIS-2	Windows (CP1250)
RIS-3	Windows (CP1250)
RIS-4	UTF-8

⁹⁵ IBM je zkratka názvu *International Business Machines*.

⁹⁶ Z prostorových důvodů je místo výrazu „identifikátor“ používána zkratka ID.

ID sady	Kódování rozpoznané editorem PSPad
RIS-5	Windows (CP1250)
RIS-6	UTF-8
RIS-7	Windows (CP1250)
RIS-8	Windows (CP1250)
RIS-9	UTF-8
RIS-10	UTF-8
RIS-11	Latin II (CP852)
RIS-12	Windows (CP1250)

Pokud se týká rozlišení začátku a konce záznamu, je ve specifikaci formátu uvedeno, že každý záznam musí začínat tagem „TY - “ a končit tagem „ER - “. ⁹⁷ Specifikace dále požaduje, aby jednotlivá pole a jejich hodnoty byly vždy umístěny na zvláštním řádku, přičemž jim musí předcházet znaky CR a LF. Z uvedeného lze vyvodit, že je primárně počítáno s užitím formátu v DOSových ⁹⁸ systémech, neboť v unixových systémech není pro ukončení řádku používán znak CR, ale pouze znak LF. Z tab. 8 je ovšem zřejmé, že použití unixových zakončení řádků je relativně časté – v testovacích sadách je to konkrétně v devíti případech z dvanácti.

Tab. 8: Oddělovací řetězce znaků ve formátu RIS

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
RIS-1	TYSS-S	LFLF	ERSS-SLFLF ⁹⁹	Unix
RIS-2	LFTYSS-S	–	ERSS-S LF	Unix
RIS-3	LFLFSLFLFTYSS-S	–	ERSS-SSLFLFLF	Unix
RIS-4	BOM ¹⁰⁰ CRLF TYSS-S	LFSSCRLF	ERSS-SCRLFSS	DOS
RIS-5	TYSS-S	CRLF CRLF	ERSS-SCRLF CRLF CRLF	DOS
RIS-6	TYSS -S	CRLF CRLF	ERSS-SCRLF	DOS

⁹⁷ Ve specifikaci je konkrétně uveden výraz "ER - ", z příkladů však plyne, že správná podoba tagu je "ER - " (čárka by správně měla být uvedena až za uvozovkami, jedná se pravděpodobně o chybu vzniklou při úpravách příslušné věty).

⁹⁸ DOS je zkratka výrazu *Disk Operating System* (diskový operační systém).

⁹⁹ Resp. nic, jsou to znaky konce záznamu, který je poslední v sadě.

¹⁰⁰ Na tomto místě byla v souboru vložena sekvence znaků EFBBBF (zapsáno hexadecimálně). Jedná se o tzv. BOM (*Byte Order Mark*, označení pořadí bytů).

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
RIS-7	HTLF LFLF LFLF SLFLF LFLF LFLF Provider: S ProQuest LLC LF Database: S ProQuest LFLF Content: S application/x-research-info-systems LFLF TYSS-S	SLFLF LFLF	ERSS-SSLFLFLF	Unix
RIS-8	HTLF LFLF LFLF SLFLF LFLF LFLF LFLF Provider: S ProQuest LLC LF Database: S ProQuest LFLF Content: S text/html LF DBVendor=ProQuest S LLC LF DBname=ProQuest SLF Text-encoding=ASCII S LFLF TYSS-S	LFHTHTLFHTLFLF	ERSS-SLFHTHTLFHTLFLF	Unix
RIS-9	TYSS-S	LF	ERSS-SLF	Unix
RIS-10	TYSS-S	LF	ERSS-SLF	Unix
RIS-11	TYSS-S	LFLF	ERSS-SLFLF	Unix
RIS-12	TYSS-S	LFLF	ERSS-SLFLF	Unix

Z tab. 8 je zřejmé, že v sedmi případech z dvanácti začínají soubory přímo prvním tagem záznamu, ve třech případech jsou předřazeny bílé znaky a ve dvou případech kombinace údajů o zdroji záznamů a bílých znaků. Z deseti sad s více záznamy se pouze ve dvou případech vyskytovalo oddělení znakem LF a v dalších dvou případech oddělení znaky CR a LF. Ve zbylých šesti případech se setkáváme s více znaky LF, popř. s kombinací dalších netisknutelných znaků. Pokud se týká posledních znaků souborů, ve všech případech se setkáváme s alespoň jedním znakem LF v případě unixových zakončení řádků, resp. CR a LF v případě DOSových zakončení řádků.

Pokud se týká způsobu práce s netisknutelnými znaky ve formátu RIS, lze ještě doplnit, že podle specifikace je možné (avšak, jak je zdůrazněno, nikoliv nutné) vložit znak konce řádku¹⁰¹ po 70 znacích a další část textu uvést až na dalším řádku. Při práci s dvanácti sadami záznamů bylo zjištěno, že v deseti případech není tento způsob rozdělování dlouhého textu využíván, v jednom případě (sada RIS-3) nebylo

¹⁰¹ Specifikace opět explicitně uvádí znaky CR a LF, jímž v unixových systémech odpovídá znak LF.

toto možné ověřit z důvodu absence dlouhého textu v polích záznamu. Pouze v sadě RIS-11 bylo rozdělování dlouhého textu zjištěno, avšak nebylo zcela v souladu se specifikací. Tuto skutečnost názorně dokládá příklad záznamu z knihy na obr. 16. V poli s šestiznakovým názvem („N1 -“) na prvním řádku po těchto šesti znacích následuje 55 znaků (hodnota pole), dále je uveden znak LF, druhý řádek pak začíná čtyřmi znaky S (mezery) a pokračuje 53 znaky vlastního textu. První řádek je tedy tvořen celkem 61 znaky, druhý řádek 57 znaky (znak LF není do tohoto počtu zahrnut).

0	10	20	30	40	50	60	70
TY - BOOK							
A1 - Li,LiLi							
T1 - Emerging technologies for academic libraries in the digital age							
Y1 - 2009							
CY - Oxford							
PB - Chandos							
SN - 1843343231							
T3 - Chandos information professional series							
N1 - 284 s. This book is written to promote academic strategic management and envision future innovations for academic library resources, services and instructions in the digital age. It provides academic executives, consultants, instructors, IT specialists, librarians, LIS students, managers, trainers and other professionals with the latest information for developing trends of emerging technologies applied to student-centred and service-oriented academic learning environments. This book explores various fields where key emerging technologies may have great implications on academic library information technologies, academic library management, academic library information services, and academic library internal operations.							
KW - Forskningsbiblioteker							
KW - Bibliotekssystemer							
KW - Academic libraries							
KW - Technological innovations							
ER -							

Obr. 16: Zobrazení záznamu knihy ze sady RIS-11

Lze tedy shrnout, že znaky CR a LF (popř. v praxi též LF v případě unixových souborů) hrají roli konce pole. Neplatilo by to pouze v případě, kdy by bylo využito možnosti rozdělit dlouhý text na více řádků – díky skutečnosti, že každé pole začíná svým názvem (tagem), po kterém následuje jeho hodnota, je tento případ snadno strojově rozpoznatelný.

V případě, že má jeden typ údaje v jednom záznamu více hodnot, je v záznamu vždy zopakován název (tag) údaje a po něm následuje jeho hodnota. Ve

specifikaci je tento způsob řešení situace explicitně uveden u autorských údajů a u klíčových slov, názorně to dokládají i příklady ve specifikaci.

Sémantika

Jak již bylo naznačeno, jednotlivé typy údajů jsou umístěny v polích s názvy (tagy), které mají podobu šesti znaků. Na první pozici je velké písmeno, na druhé velké písmeno nebo číslice, na třetí a čtvrté pozici jsou mezery, na páté pozici spojovník a na šesté opět mezera.

Údaje o autorech a editorech jsou ve formátu rozlišeny pomocí názvů polí – údaje o (primárních) autorech jsou uváděny do polí s názvy „A1 - “ nebo „AU - “, údaje o editorech (sekundárních autorech) do polí s názvy „A2 - “ nebo „ED - “.

Ve formátu jsou rozlišena i pole pro uvádění zkrácených a úplných názvů časopisů. Zatímco do polí s názvy „JF - “ nebo „JO - “ mají být uváděny úplné názvy časopisů, do pole s názvem „JA - “ pak standardní zkratky (je doporučeno je vytvářet ve stylu Index Medicus). Do polí „J1 - “ či „J2 - “ je možné zahrnout jinou (uživatelskou) zkrácenou podobu názvu časopisu.

Celkem je možné používat 49 tagů:

TY -	ED -	J1 -	M2 -
ER -	A3 -	J2 -	M3 -
ID -	Y1 -	VL -	U1 -
T1 -	PY -	IS -	U2 -
TI -	Y2 -	CP -	U3 -
CT -	N1 -	SP -	U4 -
BT -	AB -	EP -	U5 -
T2 -	N2 -	CY -	UR -
BT -	KW -	PB -	L1 -
T3 -	RP -	SN -	L2 -
A1 -	JF -	AD -	L3 -
AU -	JO -	AV -	L4 -
A2 -	JA -	M1 -	

Jak již bylo zmíněno, u některých prvků je možnost volby tagu, proto je skutečný počet polí (z hlediska obsahu) nižší.

Pravidla pro tvorbu obsahu

Pokud se týká obsahu jednotlivých polí, specifikace formátu ponechává jeho uživatelům značnou volnost. Např. u pole s tagem „AV - “ (*Availability*, tj. Dostupnost) je uveden pouze název pole a je konstatována skutečnost, že se jedná o alfanumerické pole s neomezenou délkou. Podrobněji jsou pravidla pro tvorbu obsahu rozepsána např. u již zmiňovaného pole s primárními autorskými údaji (tagy „A1 - “ nebo „AU - “) – ve specifikaci je předepsáno pořadí jednotlivých jmen (příjmení, první jméno, sufix).

Tato volnost v praxi znamená, že se můžeme setkat s různými způsoby implementace formátu RIS, resp. užití jeho polí a obsahu v nich.

Uvedené tvrzení mohou konkrétně doložit výskyty hodnot polí s autorskými údaji a údaji o názvech časopisů v testovacích datových sadách.

Tab. 9 ukazuje, že v poli s názvem „JF - “, které bylo přítomné v osmi ze zkoumaných sad, se vždy nacházel úplný název časopisu. Ve třech sadách se nacházelo pole s názvem „JA - “, z toho v případě dvou sad, v nichž se vyskytoval záznam článku z časopisu s jednoslovným názvem, který se nezkracuje, nebylo možné posoudit, zda v tomto případě obsah plnil pole funkci zkráceného či úplného názvu. V případě třetí sady bylo zjištěno uplatnění zkrácených názvů bez teček i zkrácených názvů s tečkami (rovněž se vyskytly případy uvedení úplného znění jednoslovného názvu). Pole s názvem „JO - “ se vyskytovala celkem ve čtyřech sadách, ve všech případech byly uvedeny úplné názvy časopisů. Pole s názvem „J1 - “ bylo identifikováno pouze v jedné sadě, přičemž v něm byly uvedeny úplné názvy časopisů.

Tab. 9: Pole s názvy zdrojových dokumentů ve formátu RIS

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
RIS-1	JF -	25	vždy úplný název časopisu	Signal Processing
RIS-1	JA -	25	3 výskyty úplného názvu (jednoslovný název) 4 výskyty zkráceného názvu bez teček 18 výskytů zkrácených názvů s tečkami	VINE Signal Process Environ. Model. Softw.

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
RIS-2	JA -	1	1 výskyt úplného názvu (jednoslovný název)	Nature
RIS-3	JF -	1	1 výskyt úplného názvu (jednoslovný název)	Science
RIS-3	JA -	1	1 výskyt úplného názvu (jednoslovný název)	Science
RIS-4	JF -	25	25 výskytů úplných názvů	International Journal on Digital Libraries
RIS-5	JO -	25	25 výskytů úplných názvů	Library Collections, Acquisitions, and Technical Services
RIS-6	JF -	25	25 výskytů úplných názvů	Journal of Digital Information Management
RIS-6	JO -	25	25 výskytů úplných názvů	Journal of Digital Information Management
RIS-6	J1 -	2	2 výskyty úplných názvů	Library Quarterly
RIS-7	JF -	6	6 výskytů úplných názvů	Pittsburgh Post - Gazette
RIS-7	JO -	19	19 výskytů úplných názvů	Journal of Digital Imaging
RIS-8	JF -	6	6 výskytů úplných názvů	Pittsburgh Post - Gazette
RIS-8	JO -	19	19 výskytů úplných názvů	Journal of Digital Imaging
RIS-9	JF -	1	1 výskyt úplného názvu (včetně místa vydání)	LIBRARY JOURNAL -NEW YORK-
RIS-10	JF -	1	1 výskyt úplného názvu (včetně místa vydání)	LIBRARY JOURNAL -NEW YORK-

Pokud se týká údajů o autorské odpovědnosti ve zkoumaných sadách, v sedmi případech bylo zjištěno použití názvu pole „AU -“, v pěti případech použití názvu pole „A1 -“. Počet výskytů pole v sadě se pohyboval od jednoho výskytu (sada RIS-3 obsahující jeden záznam) až po 118 výskytů (sada RIS-12 obsahující 71 záznam), celkový počet výskytů polí s autorskými údaji dosáhl 614. Z toho se v 563 případech jednalo o personální autorské údaje – v 543 případech o výskyt invertované podoby a ve 20 případech pak o výskyt neinvertované podoby.

Ve zbývajícím 51 případě se jednalo o následující údaje:

- hodnota Anonymous (20 výskytů),
- nerelevantní hodnoty (2 výskyty),
- názvy korporací – organizací (13 výskytů),
- názvy korporací – konferencí (12 výskytů),

- názvy služeb organizací (4 výskyty).

Možnost uvádět do polí pro autorské údaje jiné údaje než údaje o personální autorské odpovědnosti není ve specifikaci formátu RIS uvedena, je počítáno pouze s autory – osobami. Lze proto konstatovat, že rámcově je v souladu se specifikací formátu pouze 543 případů výskytů invertované podoby jména autora (osoby).

Vzhledem ke skutečnosti, že ve specifikaci je určeno, že mezi příjmením a prvním jménem je jako oddělovací znak použita čárka¹⁰², však nejsou požadavky dané specifikací splněny ve 472 případech z celkového počtu 543 výskytů invertované podoby (v sedmi případech z uvedených 472 navíc nebyla použita běžná mezera, ale znaky pro nedělitelnou mezeru). Ze 472 případů bylo zjištěno 339 výskytů celého jména autora (tj. příjmení a první jméno autora), ve 133 případech pak výskyt příjmení a iniciály prvního jména, popř. dalších jmen autora. Zvláštní případ tvoří data o autorech ze sad RIS-9 a RIS-10, kde je přítomna také interpunkce (čárky, tečky či jejich kombinace) na konci hodnoty pole, a to i v případech, kdy se nejedná o iniciálu jména, za níž může být logicky tečka uvedena. Zdrojem těchto dat je souborný katalog WorldCat, proto je zřejmé, že se jedná o důsledek převzetí dat z formátu MARC 21.

Požadavky jsou naopak splněny v 71 případě. Z nich je ve většině (63 výskytů) uvedeno celé jméno v invertované podobě, ve zbylých osmi případech bylo identifikováno příjmení doplněné o iniciálu prvního jména autora, popř. jeho dalších jmen. Tyto případy jsou tedy jediné, které byly zcela v souladu se specifikací formátu RIS.

Dále je nutné konstatovat, že celkem ve 36 polích s autorskými údaji bylo zjištěno nesprávné zobrazení znaků.

Podrobný rozpis zjištění týkajících se jednotlivých sad je k dispozici v tab. 10.

¹⁰² Ve specifikaci je jako povinná uvedena syntax "Lastname, Firstname, Suffix", z bezprostředně následujících příkladů (např. „Phillips,A.J“, „Phillips,Albert John“ a „Phillips,Albert“) je však zřejmé, že po čárce nemá následovat mezera.

Tab. 10: Pole s autorskými údaji ve formátu RIS

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
RIS-1	AU -	76	76 výskytů invertované podoby (mezera mezi příjmením, čárkou a iniciálami jména)	Birkinshaw, S.J. Clavel-Merrin, G. Yañez, F.
RIS-2	AU -	7	7 výskytů invertované podoby (mezera mezi příjmením, čárkou a úplným zněním jména)	Reyes, Reinabelle Smith, Robert E.
RIS-3	A1 -	1	1 výskyt invertované podoby (mezera mezi příjmením, čárkou a úplným zněním jména)	Mervis, Jeffrey
RIS-4	AU -	83	81 výskytů invertované podoby (28 výskytů: mezera mezi příjmením, čárkou a iniciálami jména; 53 výskytů: mezera mezi příjmením, čárkou a úplným zněním jména) 2 výskyty polí bez relevantních hodnot (jeden výskyt pole s čárkou, jeden výskyt prázdného pole)	Hinze, Annika Costabile, M.F. Gonçalves, Marcos Fox, Edward Mani-Saada, Jane Karadkar, Unmil P. Lin, Chia-Hung Fox, Edward A.
RIS-5	AU -	60	60 výskytů invertované podoby (56 výskytů: mezera mezi příjmením, čárkou a úplným zněním jména; čtyři výskyty: mezera mezi příjmením, čárkou a iniciálami jména)	Woodley, Mary S. Jeevan, V. K. J. Kluiters, Chris C.P. Xie, Hong Iris Fabiola Rosales, N. (Iris) Xie, Hong
RIS-6	AU -	54	52 výskytů invertované podoby (52 výskytů: mezera mezi příjmením, čárkou a úplným zněním jména; z toho čtyři výskyty: celá jména velkými písmeny) 2 výskyty neinvertované podoby	Bayram, Özlem Lai, Smyth A. Dalbello, Marija Dalbello, Marija Chew-Hung Chang John G. Hedberg Gensel, Jérôme
RIS-7	AU -	26	7 výskytů invertované podoby (bez mezery mezi příjmením a iniciálami jména) 9 výskytů neinvertované podoby (úplné znění jména) 10 výskytů hodnoty Anonymous	Welte,F. Pauline Nicholas Anonymous Mark Glover
RIS-8	AU -	26	7 výskytů invertované podoby (e znakem mezi příjmením a iniciálami jména) 9 výskytů neinvertované podoby (úplné znění jména) 10 výskytů hodnoty Anonymous	Welte, F. Pauline Nicholas Anonymous Mark Glover

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
RIS-9	A1 -	49	6 výskytů názvů organizací (na konci znaky „,“) 2 výskyty názvů služeb organizací (na konci znak „,“) 6 výskytů názvů konferencí (na konci znak „,“) 35 výskytů invertované podoby (29 výskytů: mezera mezi příjmením a jménem, na konci pole tečka, z toho ve 27 případech po úplném znění jména a ve dvou případech po iniciále; čtyři výskyty: mezera mezi příjmením a jménem, na konci pole čárka; dva výskyty: mezera mezi příjmením a iniciálou, za iniciálou, resp. na konci pole tečka) často nesprávné zobrazení znaků (chybné zobrazení v devíti polích)	Borgman, Christine L., Papy, Fabrice. Ebooks Corporation., BÄ¶rner, Katy. NÄ¡rodnÄ¡ knihovna CeskÄ© republiky., SpringerLink (Online service), DigitÄ¡lnÄ¡ knihovny, provoz a budovÄ¡nÄ¡, SedlÄ¡cek, Jan.
RIS-10	A1 -	49	6 výskytů názvů organizací (na konci znaky „,“) 2 výskyty názvů služeb organizací (na konci znak „,“) 6 výskytů názvů konferencí (na konci znak „,“) 35 výskytů invertované podoby (29 výskytů: mezera mezi příjmením a jménem, na konci pole tečka, z toho ve 27 případech po úplném znění jména a ve dvou případech po iniciále; čtyři výskyty: mezera mezi příjmením a jménem, na konci pole čárka; dva výskyty: mezera mezi příjmením a iniciálou, za iniciálou, resp. na konci pole tečka) často nesprávné zobrazení znaků (chybné zobrazení v devíti polích)	Borgman, Christine L., Papy, Fabrice. Ebooks Corporation., BÄ¶rner, Katy. NÄ¡rodnÄ¡ knihovna CeskÄ© republiky., SpringerLink (Online service), DigitÄ¡lnÄ¡ knihovny, provoz a budovÄ¡nÄ¡, SedlÄ¡cek, Jan.
RIS-11	A1 -	65	64 výskytů invertované podoby (63 výskyty: bez mezery mezi příjmením a úplným jménem; jeden výskyt: bez mezery mezi příjmením a iniciálou jména) 1 výskyt názvu organizace často nesprávné zobrazení znaků (chybné zobrazení v pěti polích)	Courtney,Nancy Witten,I. H. Nichols,David M. Borbinha,Jos b Borbinha,JosÚ Jurik,Bolette Ammitzb°ll Jurik,Bolette Ammitzb°ll Council on Library and Information Resources
RIS-12	A1 -	118	118 výskytů invertované podoby (104 výskytů: mezera mezi příjmením a úplným jménem; z toho v jednom	Pulgar Vernalte, Francisca Holt, ?lkay

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
			případě výskyt sufixu; čtrnáct výskytů: mezera mezi příjmením a iniciálou jména) často nesprávné zobrazení znaků (chybné zobrazení v třinácti polích)	Karasözen, Bülent ?????????, ?????? ?? Bailey, Jr., Charles W.

Poznámky

Existence oficiální XML verze formátu nebyla zjištěna, nepodařilo se tedy získat ani její specifikaci.

4.4.2 ISI

Specifikace

Formát ISI (také označován obecněji jako *Field Tagged Format*, tj. Formát s tagy polí) je, jak již název napovídá, používán v produktech nabízených ISI, resp. společností Thomson Reuters. Jeho specifikace (ani údaje o její aktualizaci) není veřejně přístupná. Z toho důvodu byl firmě Thomson zaslán dotaz s žádostí o zaslání specifikace formátu v elektronické podobě, popř. odkazu na ni. Součástí odpovědi však nebyla specifikace formátu, pouze vzorový záznam v tomto formátu, upřesnění podoby začátku a konce záznamu a stručné konstatování, že obdobně vypadají další záznamy [167].

Vzhledem k uvedeným skutečnostem byla většina dále uvedených údajů zjištěna analýzou datové sady s identifikátorem ISI-1. V tomto případě tedy nebylo možné porovnávat specifikaci formátu s jeho konkrétními implementacemi.

Syntax

Syntax formátu ISI si přiblížíme na příkladu datové sady ISI-1. Pokud se týká kódování, textový editor PSPad rozpoznal kódování UTF-8 (viz též tab. 11). Lze předpokládat, že toto kódování je uplatňováno i v případech dalších záznamů a jejich sad.

Tab. 11: Kódování souboru ve formátu ISI

ID sady	Kódování rozpoznané editorem PSPad
ISI-1	UTF-8

Sada záznamů začíná dvěma řádky, z nichž na prvním je uveden typ souboru (*file type*, tag „FN“¹⁰³) a číslo verze (*version number*, „VR“), a je ukončena řádkem s tagem „EF“ (lze odvodit, že se jedná o zkratku výrazu *end of file*). Jednotlivé záznamy pak začínají řádkem, na němž je uveden typ dokumentu, resp. publikace (tag PT, lze předpokládat rozpis *publication type*) a je ukončen tagem značícím konec záznamu (*end of record*, tag ER). Jednotlivé záznamy v sadě jsou od sebe navíc odděleny prázdnými řádky. Podrobný rozpis uvádění počátečních a ukončovacích tagů včetně souvisejících bílých znaků je uveden v tab. 12.

Tab. 12: Oddělovací řetězce znaků ve formátu ISI

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
ISI-1	FNSISISExportSFormatLFVRS1.0	LFLF	ERLFLFEF	Unix

Výjimku tvoří případ, kdy se v záznamu vyskytuje pole s tagem „AB“ – po hodnotě v tomto poli za znakem LF přímo tag dalšího pole, ale jsou přidány ještě tři mezery a další znak LF.

V případě, že jeden typ údaje má více hodnot, jsou údaje o druhých, popř. dalších hodnotách uváděny na nových řádcích, není však již opakováno označení údaje formou tagu. Místo tagu jsou uvedeny tři mezery, tj. první znaky hodnot polí jsou v těchto případech vždy přímo pod sebou.

Počet znaků, po nichž je v případě delšího textu v poli řádek zalomen, tj. je uveden znak LF, za nímž následují tři znaky mezery a pokračuje vlastní obsah pole, není konstantní; konstantní je však maximální počet znaků uvedený na řádce – jedná se o 74 znaků (se započtením úvodních tří mezer), resp. 71 znak (bez započtení úvodních mezer).

¹⁰³ V analyzovaných příkladech získaných ze systému Web of Knowledge byl vždy na tomto řádku uveden výraz „FN ISI Export Format“.

Znak **LF** má tedy ve formátu význam zakončení hodnoty pole nebo její přerušení z důvodu délky textu. Oba případy jsou strojově obtížně rozlišitelné, neboť výskyt znaku **LF** následovaného třemi mezerami může znamenat přerušení hodnoty pole z důvodu délky textu (typicky se jedná o případy pole s tagem „AB“, tedy o abstrakty), ale také výskyt více hodnot v poli (příkladem může být výskyt údajů o více autorech). Situaci navíc komplikuje přítomnost dalších bílých znaků (již zmíněných tří mezer a znaku **LF**) na konci hodnoty pole s tagem „AB“. Názorně to dokládá obr. č. 17.

```

PT J
TI TWO LIBRARIANS, AN ARCHIVIST, AND 13,000 IMAGES: COLLABORATING TO BUILD
A DIGITAL COLLECTION
AU Hunter, NC
   Legg, K
   Oehlerts, B
SO LIBRARY QUARTERLY
VL 80
IS 1
BP 81
EP 103
PY 2010
TC 0
AB Colorado State University Libraries has been creating digitized
collections, primarily from its Archives and Special Collections unit,
since 2000. These projects involved collaboration among Archives,
Cataloging, and Digitization; the most recent and ambitious project,
digitizing 13,000 historical images of the university dating from the
1880s into the 1930s, required closer collaboration than any previous
project. The three authors, each with a distinctive role in the
project, use this case study to illustrate and discuss in detail the
different professional and technical skills and perspectives that each
brought to the project and how they learned from each other. The
authors present lessons learned in how to effectively build successful
internal partnerships to further digitization projects.

UT WOS:000273057200005
SN 0024-2519
ER

```

Obr. 17: Záznam časopiseckého článku ze sady ISI-1

Sémantika

Jednotlivé typy údajů jsou odlišovány formou tagů – jimi jsou vždy dvouznakové (konkrétně dvoupísmenné) zkratky (viz výše zmíněné příklady tagů) doplněné o mezeru.

Pokud se týká rolí autorů a editorů, lze konstatovat, že ve formátu ISI jsou rozlišovány – zatímco u autorů se setkáváme s tagem „AU“, u editorů s tagem

„ED“. Zkrácené názvy časopisů nejsou odlišovány od názvů úplných, resp. v sadě ISI-1 se v poli s tagem „SO“ vyskytovaly pouze úplné názvy časopisů, popř. sborníků příspěvků.

Úplný seznam prvků je vzhledem k nedostupnosti specifikace obtížné určit.

V datové sadě ISI-1 bylo obsaženo celkem 21 prvků s následujícími tagy:

FN	VL	AB	CY
VR	IS	UT	CL
PT	BP	SN	ED
TI	EP	DI	
AU	PY	ER	
SO	TC	CT	

Pravidla pro tvorbu obsahu

Pravidla pro tvorbu obsahu nejsou vzhledem k nedostupnosti specifikace k dispozici, lze však usuzovat, že v praxi – v informačních produktech firmy Thomson Reuters – jsou sekundárně uplatňována pravidla pro tvorbu obsahu záznamů v databázích, na jejichž základě jsou exportována data do formátu ISI.

Vzhledem k zaměření na oblast údajů o názvech časopisů a o autorech se podrobněji podíváme právě na způsob uvádění těchto údajů v datové sadě ISI-1. Upřesněme, že počet výskytů v sadě v případě názvů časopisů či jiných zdrojových dokumentů (sborníků příspěvků, které jsou svým charakterem blízké časopisům) v tab. 13 odpovídá počtu výskytů názvů pole (v případech delších názvů jsou na začátku dalších řádků uváděny tři mezery, avšak jedná se stále jednu hodnotu či jeden výraz). Naopak v tab. 14 je uveden jak počet výskytů názvů pole, tak počet výskytů jeho hodnot.

Tab. 13: Pole s názvy zdrojových dokumentů ve formátu ISI

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
ISI-1	SO	25	vždy úplný název časopisu nebo jiného zdrojového dokumentu, zejména sborníků příspěvků – v těchto	INFORMATION PROCESSING & MANAGEMENT 2009 IEEE INTERNATIONAL SYMPOSIUM ON IT IN MEDICINE & EDUCATION, VOLS 1

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
			případech byly v rámci obsahu pole uvedeny i další doplňkové údaje o svazcích	2009 IEEE INTERNATIONAL SYMPOSIUM ON IT IN MEDICINE & EDUCATION, VOLS 1 AND 2, PROCEEDINGS

V záznamech se tedy nesetkáváme s uváděním zkratk časopisů, naopak lze říci, že v případě, že zdrojovým dokumentem není časopis, jsou v poli s tagem „SO“ uváděny i další doplňující údaje týkající se zdrojového dokumentu jako celku.

Z tab. 14 je zřejmé, že údaje o autorech a editorech jsou uváděny konzistentně, přičemž se nesetkáváme se zahrnutím úplných forem prvních a dalších jmen s výjimkou příjmení. První a další jména jsou zkracována na iniciály, za nimiž se nepíše tečky; navíc v případě více než jedné iniciály nejsou jednotlivé iniciály oddělovány mezerou. Výjimku v sadě ISI-1 tvořila hodnota „[Anon]“ naznačující fakt, že u daného příspěvku nebylo možné zjistit autora.

Tab. 14: Pole s autorskými údaji ve formátu ISI

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
ISI-1	AU	25 (počet tagů), 61 (počet hodnot)	60 výskytů invertované podoby (příjmení, čárka, mezera, iniciála prvního jména či iniciály více jmen bez mezer) 1 výskyt hodnoty „[Anon]“	Kolbe-Nusser, S Wang, MR [Anon]
ISI-1	ED	6 (počet tagů), 15 (počet hodnot)	15 výskytů invertované podoby (příjmení, čárka, mezera, iniciála prvního jména či iniciály více jmen bez mezer)	Luo, Q Mastorakis, NE

Poznámky

Existence oficiální XML verze formátu nebyla zjištěna, nepodařilo se tedy získat ani její specifikaci.

4.4.3 RefWorks

Specifikace

Specifikace formátu RefWorks je k dispozici v rámci webové prezentace citačního softwaru RefWorks [17], který nyní poskytuje společnost RefWorks-COS, součást firmy ProQuest [168]. Všechny údaje týkající se formátu (včetně praktických příkladů) jsou dostupné na jedné webové stránce. Na ní však nejsou k dispozici údaje týkající se aktualizace formátu.

Syntax

Kódování souboru se záznamem či záznamy ve formátu RefWorks by mělo podle specifikace vždy být UTF-8. Kódování datové sady REF-1 je uvedeno v tab. 15 – je možné konstatovat soulad s požadavky specifikace.

Tab. 15: Kódování souboru ve formátu RefWorks

ID sady	Kódování rozpoznané editorem PSPad
REF-1	UTF-8

Každý záznam ve formátu RefWorks by měl začínat polem s tagem „RT“. Poslední tag záznamu není ve specifikaci určen, tj. může se jednat o kterýkoliv tag s výjimkou tagu „RT“. Jednotlivé záznamy by od sebe měly být odděleny jednou prázdnou řádkou, tj. budeme-li uvažovat unixový systém, pak dvěma znaky LF, v případě DOSových systémů se pak bude jednat o sled znaků CRLF CRLF.

Z tab. 16 je zřejmé, že na začátku sady REF-1 jsou před začátkem pole s tagem „RT“ navíc uvedeny bílé znaky (pět znaků LF a jeden znak S), poslední znaky sady pak odpovídají způsobu oddělení jednotlivých záznamů v sadě.

Tab. 16: Oddělovací řetězce znaků ve formátu RefWorks

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
REF-1	LFLFSLFLFRS	–	LFLF	Unix

Z příkladů uvedených ve specifikaci vyplývá, že každé pole má být umístěno na samostatném řádku, tj. znak konce řádku v rámci záznamu odpovídá konci pole.

Tj. řádkování delších textů není řešeno přímo ve specifikaci formátu – zalomení řádků za účelem zvýšení přehlednosti je možné docílit zobrazením dat ve vhodném editoru. V testovací sadě REF-1, která obsahovala jeden záznam, nebylo možné způsob rozdělení delších textů na řádky zjistit, neboť sada neobsahovala žádné pole, jehož hodnotou by byl delší text (maximální počet znaků na řádce dosáhl počtu 33 znaků bez započítání tagu pole, 35 znaků s jeho započítáním).

Ve specifikace je explicitně uvedeno řešení, kdy mají pole s tagy „A1 “ pro primární autory a „K1 “ pro klíčová slova více hodnot – v těchto případech je vždy na další řádce zopakován tag pole a následuje další hodnota.

Sémantika

Jednotlivé typy údajů jsou označovány tříznakovými tagy, přičemž prvním znakem je vždy písmeno, druhým pak písmeno nebo číslice, třetím mezera a následně je již uváděna hodnota pole. Výjimky tvoří dva tagy – jeden pětiznakový („PMID“) a jeden šestiznakový („PMCID“).

Pokud se týká rozlišení autorů a editorů, je možné konstatovat, že ve formátu RefWorks jsou jednotlivé role velmi přesně odlišeny (k odlišení autorů jsou v pěti případech ze šesti užita adjektiva naznačující míru podílu autora na díle – od primárních autorů až po sestavovatele). Jejich přehled je uveden v tab. 17.

Tab. 17: Rozlišení různých typů autorů a editorů ve formátu RefWorks

Název pole	Vysvětlivka
A1	primární autoři
A2	sekundární autoři (editoři)
A3	terciární autoři (editoři edic)
A4	kvarternární autoři (překladaelé)
A5	kvinární autoři (sestavovatelé)
A6	editoři webových sídel

Formát RefWorks obsahuje celkem 67 prvků s následujícími tagy:

RT SR ID A1

T1	PP	DB	U12
JF	A3	DS	U13
JO	A4	IP	U14
YR	A5	RD	U15
FD	T3	ST	UL
VO	SN	U1	SL
IS	AV	U2	LL
SP	AD	U3	CR
OP	AN	U4	WT
K1	LA	U5	A6
AB	CL	U6	WV
NO	SF	U7	WP
A2	OT	U8	OL
T2	LK	U9	PMID
ED	DO	U10	PMCID
PB	CN	U11	

Formát RefWorks umožňuje rozlišení úplné a zkrácené formy názvů časopisů, a to prostřednictvím pole „JF“ pro úplný název časopisu a pole „JO“ pro zkrácený název.

Pravidla pro tvorbu obsahu

Lze konstatovat, že ve specifikaci formátu RefWorks jsou předepsána konkrétní pravidla pro vyplňování údajů, která je možné pokládat za relativně jednoznačná. Tuto skutečnost si názorně doložíme upozorněním na pravidla týkající se autorských údajů a názvů časopisů a na některá vybraná další pravidla.

Pokud se týká autorských údajů, ve specifikaci formátu je určeno, že nejprve je vždy uváděno příjmení, po něm následuje čárka, první jméno (nebo iniciála s tečkou) a iniciála středního jména s tečkou. Bezprostředně následující příklady jsou však nejednotné – příklad první hodnoty pole („Green,Adam J.,Dr.“) je v souladu se specifikací, navíc je však doplněna čárka po iniciále středního jména následovaná zkratkou titulu zakončenou tečkou. V dalších dvou příkladech („Smith, J.R.“ a „Luck, Emma“) je však po příjmení a čárce vložena ještě mezera, teprve za ní pokračují další údaje (iniciály oddělené tečkou, resp. úplné znění prvního jména). Budeme-li za správný tvar považovat tvar bez mezery, potom není hodnota údaje ze záznamu ze sady REF-1 v souladu se specifikací; budeme-li však vycházet

i z příkladů uvedených ve specifikaci, pak lze obsah pole pokládat za správně vyplněný (blíže viz tab. 18).

Tab. 18: Pole s autorskými údaji ve formátu RefWorks

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
REF-1	A1	1	1 výskyt invertované podoby (příjmení, čárka, mezera, úplné znění prvního jména)	Mervis, Jeffrey

Zkrácený název časopisu v poli „JO“ má být uváděn s tečkami (bez mezer), jak ve specifikaci ilustruje příklad „Am.Behav.Sci.“. V testovací sadě REF-1 nemohla být správnost ověřena, neboť název časopisu Science je jednoslovný, a proto se nezkracuje (viz tab. 19).

Tab. 19: Pole s názvy zdrojových dokumentů ve formátu RefWorks

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
REF-1	JF	1	1 úplný název časopisu	Science
REF-1	JO	1	1 zkrácený název časopisu (jednoslovný název shodný se zněním úplného názvu)	Science

Ve specifikaci formátu je dále mj. k dispozici taxativní výčet možných hodnot údajů v polích s tagy „RT“, „SR“, „OL“, dále jsou určeny možné atributy fontů v polích s názvovými údaji¹⁰⁴ (pole s tagy „T1“, „T2“ a „T3“), poznámkami (pole s tagem „NO“), abstrakty (pole s tagem „AB“) a uživatelskými poli 1-5 (pole s tagem „U1“, „U2“, „U3“, „U4“ a „U5“), je rovněž uveden žádoucí způsob práce s interpunkcí, uvádění zkrácených názvů časopisů (pole s tagem „JO“) a roků vydání (pole s tagem „YR“) vydání. V případě uvádění podrobnějších časových údajů v (pole s tagem „FD“) nejsou dána konkrétní pravidla pro uvádění obsahu.

Poznámky

¹⁰⁴ Určení konkrétních tagů bylo provedeno na základě charakteristiky pole z hlediska obsahu. Ve specifikaci je tedy např. uvedeno, že určité pravidlo se vztahuje k polím s názvovými údaji, není již explicitně uveden výčet tagů těchto polí.

Existuje oficiální verze formátu v jazyce XML – jedná se o jeden z nabízených exportních formátů v citačním manažeru RefWorks. Tuto verzi formátu RefWorks si přiblížíme na příkladu exportu sady 25 záznamů importovaných do citačního softwaru RefWorks z citační databáze Scopus (ukázka části prvního záznamu je k dispozici na obr. 18).

Přestože se jedná o soubor s daty v jazyce XML, bylo softwarem RefWorks implicitně nabídnuto jejich uložení ve formátu TXT. S ohledem na další práci s daty v XML editoru oXygen byla přípona souboru změněna na XML.

```
<?xml version="1.0" encoding="utf-8"?>
<refworks xmlns:refworks="www.refworks.com/xml/">

<reference>
<rt>Journal Article</rt>
<sr>Print(0)</sr>
<id>72</id>
<a1>Axelsson,I.</a1>
<a1>Marnetoft,S.-U</a1>
<t1>Benefits and harms of sick leave: Lack of randomized, controlled trials</t1>
<jf>International Journal of Rehabilitation Research</jf>
<jo>Int.J.Rehabil.Res.</jo>
<yr>2010</yr>
<vo>33</vo>
<is>1</is>
<sp>1</sp>
<op>3</op>
<k1>Randomized controlled trials</k1>
<k1>Sick leave</k1>
<k1>Sickness benefit</k1>
<k1>Campbell</k1>
<k1>Car accidents</k1>
<k1>Cochrane library</k1>
<k1>Controlled trial</k1>
<k1>Health issues</k1>
<k1>Musculoskeletal disorders</k1>
<k1>Randomized controlled trial</k1>
```

Obr. 18: Část prvního záznamu ze sady v XML verzi formátu RefWorks

Z obr. 18 je zřejmé, že po úvodní deklaraci je jako první uveden prvek „refworks“ s názvem jmenného prostoru. Tento prvek označuje celou sadu záznamů. Jednotlivé záznamy jsou pak dětmi tohoto prvku – jsou odděleny prostřednictvím prvku „reference“. Názvy jednotlivých prvků na další úrovni hierarchie vycházejí z názvů polí formátu RefWorks, avšak jsou uváděny malými písmeny a bez znaku mezery. Zatímco ve formátu RefWorks se setkáváme např. s tagem „RT“, v XML verzi formátu mu odpovídá prvek s názvem „rt“ ohraničený špičatými závorkami.

U prvků nejsou využívány atributy ani další hierarchie, jedná se tedy o velmi jednoduchou verzi XML formátu. Celkový počet prvků proto odpovídá počtu prvků ve formátu RefWorks, resp. se zahrnutím prvků „refworks“ a „reference“ se jedná o 69 tagů.

Kromě XML verze formátu RefWorks nabízí tento citační software také výstup v dalším XML formátu, který je však již označen jako neaktuální, překonaný, resp. nevhodný pro další užívání (*deprecated*). Ukázka části prvního záznamu z vytvořené sady je k dispozici na obr. 19.

```
<?xml xmlns:s='uuid:BDC6E3F0-6DA3-11d1-A2A3-00AA00C14882'  
  xmlns:dt='uuid:C2F41010-65B3-11d1-A29F-00AA00C14882'  
  xmlns:rs='urn:schemas-microsoft-com:rowset'  
  xmlns:z='#RowsetSchema'  
<s:Schema id='RowsetSchema'  
  <s:ElementType name='row' content='eltOnly' rs:CommandTimeout='30'  
    rs:updatable='true'  
    <s:AttributeType name='UserID' rs:number='1' rs:writeunknown='true'  
      rs:basecatalog='refworks' rs:basetable='refs' rs:basecolumn='UserID'  
      rs:keycolumn='true'/>  
    <s:datatype dt:type='int' dt:maxLength='4' rs:precision='10'  
      rs:fixedlength='true' rs:maybenull='false'/>  
    </s:AttributeType>  
    <s:AttributeType name='InternalRefID' rs:number='2'  
      rs:nullable='true' rs:writeunknown='true' rs:basecatalog='refworks'  
      rs:basetable='refs' rs:basecolumn='InternalRefID' rs:keycolumn='true'/>  
    <s:datatype dt:type='int' dt:maxLength='4' rs:precision='10'  
      rs:fixedlength='true'/>  
    </s:AttributeType>  
    <s:AttributeType name='Created' rs:number='3' rs:nullable='true'  
      rs:writeunknown='true' rs:basecatalog='refworks' rs:basetable='refs'  
      rs:basecolumn='Created'/>  
    <s:datatype dt:type='dateTime' rs:dbtype='timestamp'  
      dt:maxLength='16' rs:scale='3' rs:precision='23' rs:fixedlength='true'/>  
    </s:AttributeType>
```

Obr. 19: Část prvního záznamu ze sady v neaktuálním XML verzi formátu RefWorks

4.4.4 EndNote

Specifikace

Formát EndNote¹⁰⁵ slouží zejména pro import dat do stejnojmenného citačního manažeru [169], [11].¹⁰⁶

Specifikace formátu je k dispozici online v rámci uživatelského manuálu k citačnímu softwaru EndNote [18, zejm. s. 209-228, 449-463, 711-739] a dále též

¹⁰⁵ Je založen na pravidlech formátu Refer/BibIX, také bývá sám tímto výrazem označován.

¹⁰⁶ Kromě desktopové aplikace EndNote je možné pracovat také s webovou verzí EndNote Web [170], která je dostupná také v rámci platformy ISI Web of Knowledge.

v rámci nápovědy v samotném programu [11, zejm. kap. Importing Reference Data into EndNote, Reference Types a Reference Types and Fields]. Specifikace je zaměřena na vysvětlení základů formátu, jejichž znalost je nutná pro úspěšný import dat do softwaru EndNote, nechybějí však ani některé podrobnější údaje, např. seznam polí uplatňovaných v záznamech jednotlivých typů dokumentů [18, s. 711-739].

Ve výše zmíněné dokumentaci jsou k dispozici dílčí údaje vztahující se k softwaru EndNote a zprostředkovaně také ke změnám formátu v čase.

Syntax

Pokud se týká kódování, je ve specifikaci uvedeno, že citační manažer EndNote ve verzi X3 podporuje kódování UTF-8, resp. Unicode. V testovacích sadách dat bylo použité kódování určeno prostřednictvím textového editoru PSPad. Ve třech sadách bylo zjištěno kódování UTF-8, v dalších třech sadách pak kódování Windows (CP1250), viz tab. 20.

Tab. 20: Kódování souboru ve formátu EndNote

ID sady	Kódování rozpoznané editorem PSPad
END-1	UTF-8
END-2	Windows (CP1250)
END-3	Windows (CP1250)
END-4	Windows (CP1250)
END-5	UTF-8
END-6	UTF-8

Jednotlivé záznamy by podle specifikace měly od sebe být odděleny prázdnou řádkou. V záznamu by měl být nejprve uveden typ dokumentu, který daný záznam reprezentuje [18, s. 453].¹⁰⁷ V každém záznamu může být zahrnuto max. 54 polí [18, s. 226 a 450]. Poslední tag záznamu specifikace neurčuje.

Jak dokládá tab. 21, v praxi se lze setkat s uváděním i neuváděním bílých znaků na začátku záznamu; příklad datové sady END-1 dokládá, že ne vždy je v záznamu nejprve zachycen typ dokumentu a teprve následně další údaje.

¹⁰⁷ Jako příklady typů reprezentovaných dokumentů mohou posloužit „%0 Book“, „%0 Journal Article“ a „%0 Thesis“ [18, s. 225].

S výjimkou sady END-5 (sada v jazyce XML) a sady END-6 byly v sadách k oddělení záznamů užity dva prázdné řádky, nikoliv pouze jeden.

Tab. 21: Oddělovací řetězce znaků ve formátu EndNote

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
END-1	LFLFSLFLF%AS	–	LFLFLF	Unix
END-2	CRLF%0S	CRLFCRLFCRLF	CRLFCRLF	DOS
END-3	CR%0S	CRCRCR	CRCR	MAC
END-4	LF%0S	LFLFLF	LFLF	Unix
END-5	<?xml version="1.0" encoding="UTF-8"?>	–	LF	Unix
END-6	%0S	LFLF	LFLF	Unix

Je-li v záznamu přítomno více hodnot jednoho typu údaje, jsou údaje uváděny na zvláštních řádcích, přičemž je vždy zopakována i identifikace údaje. Ve specifikaci je tato skutečnost explicitně uvedena v případě autorských údajů, u dalších typů údajů se již s podobným upřesněním nesetkáváme.

Pokud se týká bílých znaků, jsou kromě oddělování záznamů a polí v záznamu užity mezi tagem pole a jeho obsahem (znak S).

Konec řádku má význam konce pole, což názorně doložily i testovací sady záznamů (v sadách END-1, END-4 a END-6 byly řádky zakončeny znaky LF, v sadě END-2 znaky CR a LF, v sadě END-3 znakem CR, v sadě END-5, tj. v souboru v XML formátu nehraje konec řádku významnou roli pro další zpracování). Způsob práce s řádkováním v případě polí s delším textem formát explicitně neřeší, uživatel si případně může upravit zobrazení (zalomení řádků) v programu či aplikaci, kterou ke zobrazení dat používá (za předpokladu, že daný program či daná aplikace toto umožňuje).

Sémantika

Jednotlivé údaje jsou identifikovány znakem „%“ a vždy jedním dalším znakem (písmena anglické abecedy, číslice 0-4 a 6-9, dalších 17 znaků, mezi něž patří např. znaky „?“ a „~“). Pak následuje mezera (znak S) a hodnota údaje.

Pokud se týká údajů o autorech a editorech, lze říci, že tyto údaje jsou uváděny v různých polích. Pole s autorskými údaji jsou předznamenána tagem „%A“, pole s údaji o editorovi či sekundárním autorovi jsou uváděna do pole s tagem „%E“. Navíc jsou k dispozici pole s tagem „%Y“, kam lze vyplnit údaje o terciárním autorovi či překladateli, a pole s tagem „%?“ určené pro podřízené autory.¹⁰⁸

Pokud se týká rozlišování zkrácených názvů časopisů od úplných, je možné použít pole „%J“ (*Journal Name*) pro úplný název a pole „%O“ (*Alternate Title*) pro název zkrácený. Praktické testy importu záznamů z externích zdrojů prokázaly, že je v praxi užíváno zejména pole „%J“, přičemž jsou v něm uváděny úplné či zkrácené názvy. V tomto směru je potřeba doplnit, že software EndNote nabízí uživatelům se seznamy termínů, jejichž prostřednictvím lze dosáhnout konzistentního užití polí [18, s. 271-292]. V našem případě se ovšem zaměřujeme na samotný formát, proto je třeba konstatovat, že pro uvádění těchto hodnot nejsou stanovena jednoznačná pravidla.

Formát obsahuje celkem 54 prvků, z nichž 52 následujících je určeno pro uživatele:

%A	%N	%0	;%\$
%B	%O	%1	;%]
%C	%P	%2	;%&
%D	%Q	%3	;%(
%E	%R	%4	;%)
%F	%S	%6	;%*
%G	%T	%7	;%+
%H	%U	%8	;%^
%I	%V	%9	;%>
%J	%W	%?	;%<
%K	%X	%@	;%[
%L	%Y	%!	;%=
%M	%Z	%#	;%~

¹⁰⁸ V případě, že se jedná o korporativní autorskou odpovědnost, má být na konci řetězce v poli přidána čárka [18].

Další dva prvky Added to Library a Last Updated jsou pak určeny pro automatické vyplňování softwarem EndNote.

Pravidla pro tvorbu obsahu

Autorské údaje je možné uvádět v invertované podobě (příjmení, jméno) i v podobě přímé, tj. neinvertované (jméno příjmení), přičemž je možné uvádět iniciály nebo úplné znění jmen. V případě autorských údajů tedy nejsou pravidla jednoznačně dána, obsah pole je možné vyplnit různými způsoby.

Ve třech datových sadách (END-2, END-3 a END-4) byla jména uvedena v přímé podobě, ve třech datových sadách (END-1, END-5 a END-6) v podobě invertované. Data byla uvedena v polích „%A“, v datové sadě END-5 s daty ve formátu XML byl použit prvek „author“ ohraničený špičatými závorkami, v datové sadě END-6 bylo navíc možné se setkat s užitím pole „%E“ pro uvedení údajů o editorech. Podrobnosti nabízí tab. 22.

Tab. 22: Pole s autorskými údaji ve formátu EndNote

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
END-1	%A	1	1 výskyt jména v invertované podobě (mezi příjmením a prvním jménem uvedeny znaky čárka a mezera)	Mervis, Jeffrey
END-2	%A	92	92 výskyty jména v přímé podobě (z toho tři výskyty iniciály prvního jména, 25 výskytů iniciály středního jména a dva výskyty kulaté závorky)	Javier López Gijón Kenneth R. Fleischmann J. Michael Moore Hong (Iris) Xie
END-3	%A	92	92 výskyty jména v přímé podobě (z toho tři výskyty iniciály prvního jména, 25 výskytů iniciály středního jména a dva výskyty kulaté závorky)	Javier López Gijón Kenneth R. Fleischmann J. Michael Moore Hong (Iris) Xie
END-4	%A	92	92 výskyty jména v přímé podobě (z toho tři výskyty iniciály prvního jména, 25 výskytů iniciály středního jména a dva výskyty kulaté závorky)	Javier López Gijón Kenneth R. Fleischmann J. Michael Moore Hong (Iris) Xie
END-5	<author>	1	1 výskyt jména v invertované podobě (mezi příjmením a prvním jménem)	Bouche, Thierry

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
			uvedeny znaky čárka a mezera)	
END-6	%A	118	118 výskytů jména v invertované podobě (mezi příjmením a prvním jménem nebo jeho iniciálou čárka a mezera; z toho 18 výskytů jmen, u nichž iniciála zastupuje první jméno, střední jméno nebo obě; v jednom z těchto 18 případů bez uvedení tečky za iniciálou)	Jacob, Carlo Elaiess, Ramadan Ghosh, T. B. Chauhan, Buddhi P
END-6	%E	15	15 výskytů jména v invertované podobě (mezi příjmením a prvním jménem nebo jeho iniciálou čárka a mezera; z toho pět výskytů jmen, u nichž iniciála zastupuje první jméno, střední jméno, popř. další jména)	Josiah, Jocelyne Sen, B. K. Murthy, T. A. V.

Pokud se týká polí s tagy „%J“ s názvy časopisů (popř. sborníků příspěvků), byl v datových sadách END-1 až END-4 zjištěn výskyt úplných názvů. V datové sadě END-5 ve formátu XML byl pro uvedení úplného názvu časopisu použit prvek „secondary_title“ ve špičatých závorkách. V datové sadě END-6 se pole s názvy časopisů nevyskytovaly. Podrobnosti jsou k dispozici v tab. 23.

Tab. 23: Pole s názvy zdrojových dokumentů ve formátu EndNote

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
END-1	%J	1	1 výskyt úplného názvu časopisu (resp. stejná podoba jako zkrácená)	Science
END-2	%J	25	25 výskytů úplných názvů časopisů nebo sborníků příspěvků	International Journal of Intelligent Systems Proceedings of the American Society for Information Science and Technology
END-3	%J	25	25 výskytů úplných názvů časopisů nebo sborníků příspěvků	International Journal of Intelligent Systems Proceedings of the American Society for Information Science and Technology
END-4	%J	25	25 výskytů úplných názvů časopisů nebo sborníků příspěvků	International Journal of Intelligent Systems Proceedings of the American Society for Information Science and Technology

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
END-5	<secoundary_title>	1	1 výskyt úplného názvu časopisu	Mathematics in Computer Science
END-6	–	–	–	–

Dále je stanoveno (implicitní) označení typů dokumentů (tj. hodnota údaje označeného tagem „%0“). V případě označení typů dokumentů jsou pravidla jednoznačně dána za předpokladu, že uživatel v softwaru EndNote neupraví typ dokumentu (resp. záznamu).¹⁰⁹

Poznámky

Přestože původní formát EndNote nevyužívá jazyk XML, je k dispozici definice typu dokumentu pro zmíněný software EndNote [172] a také jeho online (webovou) verzi EndNote Web [10].¹¹⁰

Nadále se budeme zabývat DTD pro software EndNote, neboť DTD pro online verzi EndNote je prakticky totožné, liší se pouze v drobnostech (označení odlišností je dostupné na obr. 20).

```

Top line 1 Encoding: System Line end style: DOS
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT xml (record+)>
<!ELEMENT records (record+)>
<!ELEMENT record (database?, source-app?, rec-number?, foreign-keys?, ref-type?)>
<!ELEMENT database (#PCDATA)>
<!ATTLIST database
    name CDATA #IMPLIED
    path CDATA #IMPLIED
>
<!ELEMENT source-app (#PCDATA)>
<!ATTLIST source-app
    name CDATA #IMPLIED
    version CDATA #IMPLIED
>
<!ELEMENT rec-number (#PCDATA)>
<!ELEMENT foreign-keys (key+)>
<!ELEMENT key (#PCDATA)>
<!ATTLIST key
    app CDATA #REQUIRED
    db-id CDATA #IMPLIED
>
<!ELEMENT ref-type (#PCDATA)>
<!ATTLIST ref-type
    name CDATA #REQUIRED

```

Obr. 20: Část DTD pro online verzi softwaru EndNote se zvýrazněním rozdílů oproti DTD pro desktopovou verzi softwaru

Prvku „record“ může být podřazeno celkem 64 prvků:

database	electronic-resource-num
source-app	abstract
rec-number	label
ref-type	image
contributors	caption
auth-address	notes
auth-affiliation	research-notes
titles	work-type
periodical	reviewed-item
pages	availability
volume	remote-source
number	meeting-place
issue	work-location
secondary-volume	work-extent
secondary-issue	pack-method
num-vols	size
edition	repro-ratio
section	remote-database-name
reprint-edition	remote-database-provider
reprint-status	language
keywords	urls
dates	access-date
pub-location	modified-date
publisher	custom1
orig-pub	custom2
isbn	custom3
accession-num	custom4
call-num	custom5
report-id	custom6
coden	custom7

misc1

misc3

misc2

Pěti z těchto 64 prvků mohou být podřízeny další prvky (celkem 23 prvků), viz přehled v tab. 24.

Tab. 24: Prvky podřízené prvku „record“ v XML verzi formátu EndNote

Název prvku	Název podřízených prvků
contributors	authors, secondary-authors, tertiary-authors, subsidiary-authors, translated-authors
titles	title, secondary-title, tertiary-title, alt-title, short-title, translated-title
periodical	full-title, abbr-1, abbr-2, abbr-3
dates	year, pub-dates, copyright-dates
urls	web-urls, pdf-urls, text-urls, related-urls, image-urls

U jedenácti z těchto prvků je v DTD uvedena možnost doplnění o atributy – jejich přehled je k dispozici v tab. 25.

Tab. 25: Prvky s atributy v XML verzi formátu EndNote

Název prvku	Názvy atributů
database	name, path
source-app	name, version
ref-type	name
author	corp-name, first-name, initials, last-name, middle-initial, role, salutation, suffix, title
pages	end, start
reprint-status	date, status
year	day, julian, month, year
date	day, julian, month, year
image	file, name
URL	has-ut, ppv-app, ppv-ref, ppv-ut
style	color, face, font, size

Možný výskyt stejného atributu u více než jednoho prvku lze identifikovat u atributu „name“ (čtyři prvky) a u atributů „day“, „julian“, „month“ a „year“ (dva shodně definované prvky „year“ a „date“).

Ukázka exportovaných dat z citačního manažeru EndNote v tomto formátu je k dispozici na obr. 21.

```
<?xml version="1.0" encoding="UTF-8" ?>
<xml>
  <records>
    <record>
      <database name="My EndNote Library.enl"
        path="C:\Documents and Settings\Linda\My Documents\My EndNote Library.enl">My
        EndNote Library.enl</database>
      <source-app name="EndNote" version="13.0">EndNote</source-app>
      <rec-number>21</rec-number>
      <foreign-keys>
        <key app="EN" db-id="p5dxe9rs9d0t2ke5esyv2s0r0sa9tzw2z5v">21</key>
      </foreign-keys>
      <ref-type name="Journal Article">17</ref-type>
      <contributors>
        <authors>
          <author>
            <style face="normal" font="default" size="100%">Schuldt, K.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Esser, C.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Evans, J.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">May, J.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Timmann, C.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Ehmen, C.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Loag, W.</style>
          </author>
          <author>
            <style face="normal" font="default" size="100%">Ansong, D.</style>
          </author>
        </authors>
      </contributors>
    </record>
  </records>
</xml>
```

Obr. 21: Část záznamu ze sady v exportním XML formátu softwaru EndNote

Právě na tomto příkladu lze názorně doložit určité využití hierarchie (prvek „author“ je dítětem prvku „authors“ a ten ještě obecnějšího prvku „contributors“). Je využita i možnost uvádění atributů prvků, jedná se však převážně o údaje mající vztah k zobrazení dokumentu (jsou uvedeny mj. atributy „face“, „font“ a „size“). Oproti DTD je v souboru navíc uveden prvek „xml“ jako rodič prvku „records“ – právě tento prvek by měl být součástí XML exportu z online verze citačního softwaru EndNote, je tedy pravděpodobné, že ve skutečnosti bylo použito toto DTD i v desktopové verzi manažeru EndNote.

Podle specifikace jazyka XML [1], [2] je ovšem používání prvků a atributů začínajících řetězcem xml, XML či jinou kombinací velkých a malých písmem

v tomto řetězci vyhrazeno pro potřeby konsorcia W3C. Proto není exportovaný záznam s prvkem „xml“ dobře naformátovaný (*well-formed*). Většina parserů to však nerozpozná, neboť kvůli optimalizaci rychlosti načítání kódu tuto podmínku nekontroluje. Jedná se o ukázkou případu, kdy je XML formát implementován bez dostatečné znalosti jazyka XML a jeho zákonitostí.

4.4.5 BibTeX

Specifikace

Podoba formátu BibTeX je popsána v řadě zdrojů, které vycházejí z původního návrhu formátu z 80. let 20. století, za jehož vývojem stál počítačový vědec O. Patashnik [173], [174]. Zároveň je jako spoluautor uváděn počítačový vědec L. Lamport, který se stal iniciátorem vývoje typografického systému LaTeX – právě s ním je často formát BibTeX využíván a pro jeho potřeby byl původně vyvinut.

Různě podrobné údaje o formátu BibTeX lze získat z řady zdrojů, často manuálů, dokumentace či podrobných encyklopedických hesel týkající se BibTeXu či obecně LaTeXu (viz např. [175], [19], [176] či [177]).

Pro účely této práce je za specifikaci pokládán manuál N. Markeyho *Tame the Beast : The B to X of BibTeX* [19]. Tento dokument byl jako klíčový zdroj vybrán z toho důvodu, že obsahuje velmi podrobnou charakteristiku formátu BibTeX a rovněž řadu praktických doporučení týkajících se jeho využívání. Doplnkovým zdrojem je původní specifikace formátu od O. Patashnika [174]. V této podkapitole jsou v případech, kdy bylo nutné doplnit údaje z dalších zdrojů, tyto zdroje citovány v textu.

Pokud se týká aktualizace formátu, lze uvést, že formát je velmi stabilní, nedochází k jeho významným změnám v čase. Toto tvrzení nedávno názorně doložil v roce 2003 sám O. Patashnik, když uvedl, že standardní styly formátu BibTeX z roku 1988 jsou stále aktuální [178, s. 27].

Syntax

Formát BibTeX zatím nepoužívá kódování UTF-8 (či Unicode), neboť za stále aktuální lze pokládat verzi formátu z roku 1988 – v té době standard Unicode začal teprve vznikat [179]. Formát BibTeX je naopak uzpůsoben pro osmibitové kódování, tedy např. kódování ISO-8859-1 (tj. Latin-1) [180]. O zahrnutí kódování Unicode je uvažováno až pro plánovanou novou verzi formátu BibTeX, která by měla nést označení 1.0 [178, s. 28].

Pro zobrazení znaků s diakritikou či jiných znaků je však možné použít speciální syntaktické prostředky (kombinaci složených závorek, zpětného lomítka, diakritického znaménka a příslušného znaku bez diakritiky).

U zkoumaných datových sad BIB-1 až BIB-7 nicméně v šesti případech ze sedmi textový editor PSPad rozpoznal kódování jako UTF-8, viz tab. 26.

Tab. 26: Kódování souborů ve formátu BibTeX

ID sady	Kódování rozpoznané editorem PSPad
BIB-1	UTF-8
BIB-2	UTF-8
BIB-3	UTF-8
BIB-4	Windows (CP1250)
BIB-5	UTF-8
BIB-6	UTF-8
BIB-7	UTF-8

Začátek záznamu je ve formátu BibTeX určen znakem „@“. Po tomto znaku následuje typ dokumentu a vlastní záznam (s uvedením interního klíče, tj. identifikátoru) uzavřený do složených nebo kulatých závorek. Poslední závorka tak zároveň ukončuje celý záznam.

Mezery a konce řádků ve formátu BibTeX nehrají důležitou roli, jsou používány pouze k zajištění lepší čitelnosti záznamů pro uživatele – člověka.

Pokud se týká prvních znaků souborů se sadami záznamů ve formátu BibTeX, lze konstatovat, že ve dvou případech (sady BIB-4 a BIB-7) soubory začínaly přímo znakem „@“, tj. prvním znakem záznamu. V jednom případě (sada BIB-6) tomuto znaku předcházela sekvence znaků BOM, tj. prokazatelný důkaz, že

soubor je skutečně v kódování UTF-8. V případě sady BIB-3 se vyskytla kombinace znaků **LF** a **S**, v případě sada BIB-2 byla dvěma znakům **LF** předřazena úvodní data týkající se zdroje a data exportu. Sada BIB-5 byla ve skutečnosti souborem v jazyce XML, resp. HTML s rozsáhlým doprovodným kódem. Poslední znaky souborů (po konci posledního záznamu, tj. po znaku „}“), ve třech sadách (BIB-1, BIB-2 a BIB-3) se vyskytl vždy jeden znak **LF**, ve dvou sadách (BIB-4 a BIB-6) kombinace znaků **CR** a **LF** (v sadě BIB-4 se kombinace nalézala dvakrát) a v jedné sadě (BIB-7) kombinace dvou znaků **LF**. Pro úplnost dodejme, že tato kombinace se vyskytla i v sadě BIB-5, avšak vzhledem k přítomnosti dalšího kódu se jedná o specifický případ. Způsob oddělení záznamů v sadě bylo možné zjistit u sad BIB-2 a BIB-7 (vždy dva znaky **LF**) a také u sady BIB-4 (kombinace znaků **CR** a **LF**). Pokud se týká způsobu zakončení řádku, lze konstatovat, že textový editor PSPad jej v případě sady BIB-7 určil nesprávně. V souboru se totiž vyskytoval jeden znak **CR**, standardní zakončení řádku bylo však realizováno bez tohoto znaku, proto se jedná o unixové zakončení. Podrobnosti nabízí tab. 27.

Tab. 27: Oddělovací řetězce znaků ve formátu BibTeX

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
BIB-1	LF @	–	}, LF ; LF	Unix
BIB-2	Scopus LF EXPOR T SDATE:15 S Mar ch2010 LF LF @	LFLF	}, LF ; LF	Unix
BIB-3	LFLFSLFLF @	–	}; LF ; LF	Unix
BIB-4	@	CRLF	CRLF ; CRLF CRLF	DOS
BIB-5	<?xml S version="1.0" S encoding="utf-8" S ?> a následný rozsáhlý úsek kódu	<br S >	</body> CRLF </ht ml> CRLF	DOS
BIB-6	BOM@	–	, CRLF ; CRLF	DOS
BIB-7	@	LFLF	}; LF ; LFLF	Unix

Jednotlivá pole jsou od sebe ve formátu BibTeX oddělena čárkami. Po názvu polí následuje znak „=" a hodnota pole uzavřená do uvozovek (znaky „"“)

nebo složených závorek (znaky „{}“). Právě s ohledem na čitelnost jsou mezi názvem pole, rovnítkem a obsahem pole často doplňovány mezery.

Pokud se týká řádkování v polích s delším textem, ze specifikace lze usuzovat, že je možné použít právě konců řádků. Jejich užití v praxi bylo prokázáno v datových sadách BIB-1, kde byl znak LF navíc – pravděpodobně pro zvýšení přehlednosti zobrazení hodnot příslušných polí – doplněn třemi znaky S, a BIB-7, v níž byl použit pouze znak LF).

Je však zároveň třeba konstatovat, že implicitně je délka pole omezena na tisíc znaků,¹¹¹ omezení je možné překonat (např. s využitím návodu zveřejněného v rámci často kladených dotazů The UK TeX Archive [181]).

Ve specifikaci formátu BibTeX se setkáváme s konkrétními výskyty více hodnot jednoho typu údaje u obsahu pole „author“, v tomto případě jsou jednotlivé hodnoty odděleny výrazem „and“. Jsou také přesně dána syntaktická pravidla, která je nutné dodržovat, aby bylo možné z formátu generovat žádoucí výstupy. V případě formátu BibTeX tedy do značné míry dochází k prolnutí syntaktických pravidel s pravidly pro tvorbu obsahu.

Sémantika

Standardizované názvy polí mají ve formátu BibTeX podobu celých výrazů (typicky původně jednoslovných anglických). Dokládá to i výčet celkem 23 prvků obsažených ve čtyřech stylech,¹¹² jejichž autorem je O. Patashnik (jedná se o styly plain.bst, alpha.bst, unsrt.bst a abbrv.bst):

address	howpublished	organization
author	institution	pages
booktitle	journal	publisher
chapter	key	school
crossref	month	series
edition	note	title
editor	number	type

¹¹¹ Omezení je má pochopitelně význam při využití nástrojů přímo určených pro práci s formátem BibTeX, které např. upozorňují uživatele na chyby v sintaxi apod.

¹¹² Prostřednictvím stylu jsou vymezena dostupné typy záznamů, pole, která je v nich možné použít, způsob extrakce a uspořádání jednotlivých záznamů a výsledná podoba seznamu záznamů [19, s. 13].

volume year

Vzhledem ke skutečnosti, že je možné vytvářet další styly,¹¹³ popř. jednotliví uživatelé si mohou podle potřeby přidávat další pole (ta jsou při strojovém zpracování nástroji určenými pro práci s těmito daty ignorována, resp. nezobrazují se na výstupu), nejedná se o konečný seznam.

Pro zajímavost lze uvést, že byly vytvořeny i styly umožňující zachycování dat v podobě vhodné pro generování záznamů v (alespoň částečném) souladu s normami ČSN ISO 690 a ČSN ISO 690-2 (viz [182] a [183]).

Omezíme-li se na tyto čtyři standardní styly, je možné konstatovat, že údaje o autorech a editorech jsou odlišovány – jsou k dispozici pole „author“ a „editor“. Jak však upozorňuje O. Patashnik [174, s. 14], s autorskými údaji se lze v případě záznamů manuálů (typ dokumentu „manual“) či konferenčních sborníků (typ dokumentu „proceedings“) setkat v poli „organization“ (jedná se tedy o korporativní autorství).

Pokud se týká názvů časopisů, v poli „journal“, lze spíše předpokládat uvádění jejich úplných názvů s tím, že je možné použít styl `abbrv.bst`, který kromě prvních jmen autorů a názvů měsíců zkracuje také názvy předdefinovaných časopisů. Implicitně jsou však dostupné pouze zkrácené názvy vybraných časopisů z oblasti počítačové vědy [19, s. 16].

Pravidla pro tvorbu obsahu

Pravidla pro tvorbu obsahu jsou předepsána formou přesné syntaxe – jedná se zejména o pole „author“. Obecně lze konstatovat, že ve formátu jsou rozlišeny tři základní možné způsoby uvádění autorských údajů, jimiž jsou:

- "First von Last",
- "von Last, First"
- a "von Last, Jr, First".

Konkrétní způsob rozpoznání a interpretace těchto údajů je k dispozici např. v [19, s. 23-25].

V rámci čtyř „klasických“ stylů je také k dispozici výčet čtrnácti možných hodnot typů dokumentů:

¹¹³ Podrobnější přiblížení čtyř „klasických“ a vybraných dalších stylů je dostupné v [19, s. 15-17].

@article	@incollection	@phdthesis
@book	@inproceedings	@proceedings
@booklet	@manual	@techreport
@conference	@masterthesis	@unpublished
@inbook	@misc	

Pokud se týká názvů časopisů v datových sadách BIB-1 až BIB-7, lze konstatovat, že ve všech sadách je pro uvedení tohoto údaje použito pole „Journal = “ či „journal = “ (z hlediska formátu BibTeX se jedná o totožné označení, rovněž přítomnost či naopak nepřítomnost mezer je v pořádku), v sadě BIB-1 se navíc setkáme s polem „Journal-ISO = “ a v sadě BIB-2 s poli „abbreviated_source_title=“ a „coden=“ (poslední jmenované pole je ze své podstaty na pomezí pole s názvovými údaji a pole s údaji o identifikátorech zdrojového dokumentu). Tato pole již nejsou součástí čtyřech „klasických“ stylů. Obsahem pole „Journal = “ byl ve zkoumaných datových sadách vždy úplný název časopisu, resp. v případě sady BIB-3, kdy se jednalo o jednoslovný název časopisu, nebylo možné úplný od případného zkráceného názvu odlišit. V sadách BIB-4 a BIB-6 bylo zvoleno uvedení hodnoty pole v uvozovkách, ve zbývajících sadách byla dána přednost uzavření hodnoty do složených závorek. V sadě BIB-1 byly tyto závorky zdvojeny, tím je zabráněno případné změně velikosti písmen v některých stylech. Ve dvou sadách (BIB-1 a BIB-7) se vyskytly znaky, jimž by mělo předcházet zpětné lomítko, s uvedením tohoto lomítka; naopak v sadách BIB-4 a BIB-5 se přes přítomnost příslušných znaků zpětné lomítko nevyskytovalo. Podrobná zjištění jsou k dispozici v tab. 28.

Tab. 28: Pole s názvy zdrojových dokumentů ve formátu BibTeX

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
BIB-1	Journal =	1	1 výskyt úplného názvu (všechna písmena velká; užití zpětného lomítka před znakem „&“)	{{INFORMATION PROCESSING \& MANAGEMENT}},
BIB-1	Journal-ISO =	1	1 výskyt zkráceného názvu	{{Inf. Process. Manage.}},
BIB-2	journal=	25	25 výskytů úplných názvů	{Environmental Modelling and Software},

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
				{OCLC Systems and Services},
BIB-2	abbrev_source_title=	25	25 výskytů zkrácených názvů (z toho 18 výskytů zkrácených názvů s tečkami, čtyři výskyty bez teček a tři výskyty nezkrácených jednoslovných názvů)	{OCLC Syst. Serv.}, {J Power Sources}, {VINE},
BIB-2	coden=	7	7 výskytů zkrácených názvů, resp. pětimístných identifikátorů CODEN	{EMSOF},
BIB-3	journal =	1	1 výskyt úplného názvu (úplný jednoslovného název je shodný se zkráceným)	{Science},
BIB-4	journal =	25	25 výskytů úplných názvů (z toho v patnácti případech název obsahoval znak „&“ bez předcházejícího zpětného lomítka)	"The International Information & Library Review", "Library Collections, Acquisitions, and Technical Services",
BIB-5	Journal =	24	24 výskytů úplných názvů (z toho v šesti případech název obsahoval znak „&“ bez předcházejícího zpětného lomítka)	{Journal of Digital Information Management}, {International Information & Library Review},
BIB-6	journal =	1	1 výskyt úplného názvu	"Mathematics in Computer Science",
BIB-7	journal =	26	26 výskytů úplných názvů (z toho ve čtyřech záznamech výskyt znaků doplněných o zpětná lomítka)	{Open and Libraries Class Journal}, {Software libre para servicios de informaci'on digital},

Autorské údaje byly v datových sadách BIB-1 až BIB-7 uváděny jak v invertované, tak v přímé podobě. Ve čtyřech sadách (BIB-1, BIB-2, BIB-3 a BIB-5) se vyskytovaly pouze invertované podoby jmen, ve třech sadách (BIB-4, BIB-6

a BIB-7) pouze přímé podoby jmen. V pěti sadách (BIB-1, BIB-2, BIB-3, BIB-5 a BIB-7) byly hodnoty polí uzavřeny do složených závorek, ve zbývajících dvou sadách (BIB-4 a BIB-6) pak do uvozovek. V celkem pěti sadách se vyskytovala alespoň jedna hodnota pole s více než jedním autorem a bylo tedy možné identifikovat způsob oddělení jednotlivých jmen. Ve čtyřech sadách (BIB-1, BIB-4, BIB-5 a BIB-7) byla jednotlivá jména oddělena mezerou, výrazem „and“ a další mezerou, v jedné sadě (BIB-2) byla jednotlivá jména oddělena čárkou a mezerou. V této sadě se navíc na konci údajů o konkrétním autorovi v některých záznamech nalézaly dvouznakové výrazy složené z malého písmena a mezery. Ty však byly redundantní, neboť se jednalo o znaky, které jsou v rozhraní daného systému (Scopus) použity v podobě horního indexu a slouží jako odkaz na údaje o pracovišti autora. Ve všech sadách byly autorské údaje uvedeny v poli „author =“ (resp. „author=“ či „Author =“), v sadě BIB-7 navíc byly přítomné výskyty pole „editor =“ V sadách BIB-6 a BIB-7 bylo možné zjistit přítomnost jmen napsaných řeckou abecedou – příklady jsou uvedeny v tab. 29, kde jsou také k dispozici podrobnosti týkající se výskytu polí s autorskými údaji.

Tab. 29: Pole s autorskými údaji ve formátu BibTeX

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
BIB-1	Author =	1	1 výskyt invertované podoby (mezi příjmením a iniciálou prvního jména čárka a mezera; oddělení jednotlivých celých jmen mezerou, výrazem „and“ a mezerou)	{Berendt, B. and Krause, B. and Kolbe-Nusser, S.},
BIB-2	author=	25	25 výskytů invertované podoby (mezi příjmením a iniciálou prvního jména čárka a mezera, případná další iniciála následuje ihned po tečce po první iniciále a je rovněž zakončena tečkou; oddělení jednotlivých celých jmen čárkou a mezerou; v pěti záznamech jsou navíc za iniciálami uvedena malá písmena, která jsou v záznamech v databázi zobrazena formou horního indexu a slouží jako odkaz na údaje o pracovišti autora)	{Birkinshaw, S.J., James, P., Ewen, J.}, {Yañez, F.a , Chianella, I.b , Piletsky, S.A.b , Concheiro, A.a , Alvarez-Lorenzo, C.a },
BIB-3	author =	1	1 výskyt invertované podoby (mezi příjmením a úplným zněním prvního jména čárka a mezera)	{Mervis, Jeffrey},

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
BIB-4	author =	25	25 výskytů přímé podoby (jména jsou uváděna v úplné podobě, iniciály se vyskytují pouze u středních jmen, u prvního jména pouze v jednom případě; oddělení jednotlivých celých jmen mezerou, výrazem „and“ a mezerou; ve čtyřech záznamech chybně zobrazené znaky, v jednom záznamu výskyt kulatých závorek)	"Daniel G. Dorner", "V. K. J. Jeevan", "N. Fabiola Rosales and Marlene Bauste and Fernando RodrÍguez and Francisco Le'n", "Dietmar Wolfram and Hong (Iris) Xie",
BIB-5	Author =	25	25 výskytů invertované podoby (jména jsou uváděna v úplné podobě, iniciály se vyskytují pouze u středních jmen, a to celkem v pěti případech; ve dvou případech jsou autorské údaje uvedeny velkými písmeny; v jednom případě – viz poslední příklad – na první pohled podle interpunkce výskyt invertované podoby, ve skutečnosti se jedná o záměnu příjmení a prvního jména, resp. prvního jména a iniciály středního jména)	{Arif, Muhammad and Kanwal, Saima}, {Chew-Hung, Chang and John G., Hedberg}, {WOOLDRIDGE, BROOKE and TAYLOR, LAURIE and SULLIVAN, MARK}, {Chew-Hung, Chang and John G., Hedberg},
BIB-6	author =	1	1 výskyt invertované podoby (mezi příjmením a prvním jménem čárka a mezera)	"Bouche, Thierry",
BIB-7	author =	71	71 výskytů přímé podoby jména (jména jsou uváděna převážně v úplné podobě; v jedenácti záznamech se setkáváme s iniciálami, z toho se v osmi záznamech vyskytuje alespoň jedna iniciála prvního jména, v deseti případech alespoň jedna iniciála středního jména; celkem v 19 záznamech je uveden alespoň jeden znak doplněný o upřesnění znaku, zejména diakritiky, formou zpětného lomítka, ovšem bez uzavření do složených závorek; z toho ve třech případech je celý obsah pole – znaky řecké abecedy – uveden v kódované podobě doplněné o znaky „\$“ předcházející zpětnému lomítku; ve dvou z těchto záznamů jsou nesprávně rozpoznané znaky nahrazeny znakem „?“)	{Ramadan Elaie\$}, {K. Rajasekharan and K. M. Nafala}, {Celso Gonz\ales-Cam}, {\Sigma\$\alpha\$ \$\varrho\$?\$\nu\$ \$\tau\$}\$o\$ \$\varsigma\$ \$\mathrm K\$ \$\alpha\$ \$\pi\$ \$\iota\$ \$\delta\$? \$\kappa\$ \$\eta\$ \$\varsigma\$ and \$\mathrm M\$ \$\iota\$ \$\chi\$? \$\lambda\$ \$\eta\$ \$\varsigma\$ \$\Sigma\$ \$\varphi\$ \$\alpha\$ \$\kappa\$? \$\kappa\$ \$\eta\$}

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
				$\{\text{varsigma}\}$,
BIB-7	editor =	10	10 výskytů přímé podoby jména (jména jsou uváděna v šesti případech v úplné podobě, ve čtyřech záznamech se vyskytují alespoň dvě iniciály; ve třech záznamech je uveden alespoň jeden znak doplněný o zpětné lomítko bez uzavření do složených závorek, z toho se ve dvou případech jedná o znaky řecké abecedy, v obou jsou nesprávně rozpoznané znaky nahrazeny znakem „?” ¹⁴ , který se navíc nachází i v jednom dalším záznamu ze sady)	$\{\text{Anup Kumar Das}\}$, $\{\text{T. A. V. Murthy}\}$, $\{\text{Jes\us Tramullas and Piedad Garrido}\}$, $\{\text{\Gamma}\}$ $\{\text{\iota}\}$ $\{\text{\nu}\}$ $\{\text{\eta}\}$ $\{\text{\varsigma}\}$ $\{\text{\mathrm{T}}\}$ $\{\text{\sigma}\}$ $\{\text{\kappa}\}$ $\{\text{\omega}\}$ $\{\text{\nu}\}$ $\{\text{\alpha}\}$ $\{\text{varsigma}\}$, $\{\text{Barbara Szczepa? ska}\}$,

Poznámky

Existuje několik verzí formátu BibTeX v jazyce XML – v tomto textu se budeme zabývat verzí, která je dostupná prostřednictvím webové prezentace formátu [185].¹⁴ K dispozici je formát jako XML schéma (přípona .xsd), DTD (přípona .dtd) i jako kompaktní schéma Relax NG (přípona .rnc). Následující popis XML verze formátu BibTeX byl připraven na základě kompaktního schématu Relax NG, a to nejprve základního schématu (tj. souboru bibtexml.rnc) a následně rozšířeného schématu (tj. souboru bibtexml-ext.rnc).

Základní schéma obsahuje celkem 40 prvků:

entry	inproceedings	techreport
file	conference	unpublished
article	manual	address
book	masterthesis	author
booklet	phdthesis	booktitle
inbook	misc	chapter
incollection	proceedings	edition

¹⁴ Z dalších aktivit, jejichž cílem bylo zpracovat XML verzi formátu BibTeX, lze jmenovat diplomový projekt B. Luratiho a L. Previtaliho [184] či vyjádření formátu BibTeX v jazyce OWL [186].

editor	organization	volume
howpublished	pages	year
institution	publisher	key
journal	school	crossref
month	series	annote
note	title	
number	type	

Z tohoto seznamu odpovídá 23 položek prvkům ze čtyř „klasických“ stylů formátu BibTeX (viz výše), 14 typům dokumentů, s nimiž se setkáváme ve formátu BibTeX, navíc jsou obsaženy prvky „entry“, „file“ a „annote“.

Rozšířené schéma obsahuje prvky, které nejsou součástí klasického formátu BibTeX. V rozšířeném je obsažen i odkaz na základní schéma (soubor bibtexml.rnc), jsou zmíněny prvky „author“, „editor“, „title“ a „chapter“, u nichž z důvodu rozšíření schématu o prvky „person“ a „subtitle“ došlo ke změnám ve vzájemných vztazích těchto prvků.

Rozšířené schéma obsahuje celkem 25 prvků (se započítáním čtyř výše zmíněných prvků z odkazovaného základního schématu se jedná o 29 prvků):

abstract	mrnumber	initials
affiliation	price	last
contents	size	middle
copyright	url	prelast
isbn	keywords	lineage
issn	category	subtitle
language	doi	keyword
lccn	person	
location	first	

U prvku person jsou definovány dva možné atributy, a to „email“ a „url“.

Výhodou obou schémat je také skutečnost, že formou obsáhlých komentářů jsou explicitně uvedena pravidla pro tvorbu obsahu.¹¹⁵

¹¹⁵ Dění týkající se rozvoje XML verze formátu BibTeX je možné sledovat prostřednictvím RSS kanálu, bohužel poslední novinky jsou již poněkud staršího data (konkrétně z roku 2007).

4.4.6 OpenURL COinS

Specifikace

Specifikace formátu OpenURL COinS a další související informace jsou dostupné v rámci oficiální webové prezentace formátu [187]. Podstatnou roli hraje standard Z39.88-2004 týkající se OpenURL, na jehož základě byl formát OpenURL COinS vytvořen – úplné znění tohoto standardu je dostupné v rámci webové prezentace NISO [188].

Pokud se týká aktualizace, je na domácí stránce webu OpenURL COinS uvedeno, že se jedná o specifikaci stabilní verze 1.0, v poznámkách na téže stránce lze zjistit změny, které byly na webu provedeny, a data jejich realizace.

Syntax

Pokud se týká kódování, lze vycházet z normy ANSI/NISO Z39.88-2004 [188, zejm. s. 21-22]. Standard podporuje řadu různých kódování, musí se však vždy jednat o kódování, které je registrováno. V rámci registru jsou v současné době zahrnuta čtyři různá kódování (v hranatých závorkách jsou doplněny odkazy na dokumenty související s registrací u IANA či na příslušné internetové normy RFC):

- Big5-HKSCS5 [189],
- Big5 [190],
- ISO Latin 1 [191]
- a UTF-8 encoded Unicode [192].

V datové sadě OPE-1 nebylo zvolené kódování explicitně uvedeno, editor PSPad jej rozpoznal jako UTF-8 (viz tab. 30).

Tab. 30: Kódování souboru ve formátu OpenURL COinS

ID sady	Kódování rozpoznané editorem PSPad
OPE-1	UTF-8

Pokud se týká začátku a konce záznamu, lze jej podle specifikace rozpoznat podle prvku „span“, resp. počátečního tagu „span“ (s atributem „class“, který má hodnotu Z3988, a „title“) a koncového tagu „span“.

Konec řádku ve formátu s ohledem na skutečnost, že data jsou uvedena souvisle, význam nemá, není ani potřeba řešit řádkování v případě delších textů. Pokud se týká bílých znaků, je ve specifikaci (resp. v průvodci pro implementaci OpenURL 1.0 Context Object pro časopisecké články [193]) explicitně zmíněna mezera, která by měla být zakódována s využitím znaku „%“, tj. konkrétně jako „%20“. Je zmíněno, že také ostatní hodnoty mají být zakódovány s využitím tohoto znaku. Jedná se o standardní způsob zahrnutí znaků, které nejsou součástí znakové sady US-ASCII¹¹⁶ ani do této znakové sady převoditelné, do URL, resp. obecně do URI [194, s. 12].

Řešení případu více hodnot jednoho typu údaje záleží na výběru metadatového schématu (viz dále).

Oddělovací řetězce znaků jsou uvedeny v tab. 31.

Tab. 31: Oddělovací řetězce znaků ve formátu OpenURL COinS

ID sady	První znaky sady	Oddělení záznamů v sadě	Poslední znaky sady	Zakončení řádku
OPE-1	?url_ver=Z39.88-2004	–	rft.language=en	DOS

Sémantika

Ve formátu je používána značka „span“ s atributy „class“ a „title“, které jsou součástí jazyka HTML. Další údaje jsou pak hodnotami těchto atributů.

Sémantika formátu závisí na vybrané sadě metadatových prvků – jejich seznam je k dispozici v [195]. Je tvořen celkem 26 položkami, z nichž čtyři mají status „Deprecated“ a jedna má status „Review“.

Ve specifikaci formátu jsou uvedeny dva průvodce pro implementaci – jeden z nich se týká knih a druhý časopiseckých článků.

U příkladu záznamu knihy se setkáme s použitím metadatového formátu „info:ofi/fmt:kev:mtx:book“. Ten obsahuje celkem 24 různých prvků:

rft.btitle	rft.aulast	rft.auinit
rft.isbn	rft.aufirst	rft.auinit1

¹¹⁶ ASCII je zkratka výrazu *American Standard Code for Information Interchange* (Americký standardní kód pro výměnu informací).

rft.auinitm	rft.place	rft.spage
rft.ausuffix	rft.pub	rft.epage
rft.au	rft.date	rft.pages
rft.aucorp	rft.edition	rft.issn
rft.atitle	rft.tpages	bici
rft.title	rft.series	rft.genre

Pro uvádění autorských údajů jsou určeny prvky rft.aulast až rft.aucorp z výše prezentovaného výčtu (zvláštní pole pro uvádění údajů o editorech není zahrnuto). U pole rft.au, je explicitně uvedeno, že se jedná o opakovatelné pole.

U příkladu záznamu časopisu je použit metadatový formát „info:ofi/fmt:kev:mtx:journal“ obsahující celkem 29 prvků:

rft.atitle	rft.artnum	rft.aucorp
rft.title	rft.issn	rft.isbn
rft.jtitle	rft.eissn	rft.coden
rft.stitle	rft.aulast	rft.sici
rft.date	rft.aufirst	rft.genre
rft.volume	rft.auinit	rft.chron
rft.issue	rft.auinit1	rft.ssn
rft.spage	rft.auinitm	rft.quarter
rft.epage	rft.ausuffix	rft.part
rft.pages	rft.au	

Pro autorské údaje jsou používány stejné prvky jako v předchozím případě. Z našeho hlediska je potřeba zmínit ještě prvky „rft.jtitle“ (s ohledem na zpětnou kompatibilitu je možné použít také název „rft.title“) pro uvádění úplného názvu časopisu a „rft.stitle“ pro uvádění jeho zkráceného názvu.

Pravidla pro tvorbu obsahu

Pravidla pro tvorbu obsahu rovněž závisí na vybrané metadatové sadě a existenci či neexistenci pravidel či alespoň doporučení pro uvádění hodnot prvků.

U výše zmíněných příkladů se můžeme setkat s relativně podrobným vysvětlením způsobu uvádění jednotlivých údajů (často jsou k dispozici i konkrétní příklady). Např. u pole „rft.au“ je možné uvádět úplné jméno autora, ať již v přímé či

invertované podobě, do polí „rft.aulast“, „rft.aufirst“, „rft.auinit“, „rft.auinit1“ a „rft.ausuffix“ pak jednotlivé části jména, jak naznačují zkrácené podoby anglických slov, do pole „rft.autcorp“ je možné uvést název korporace, je-li autorem dokumentu.

V datové sadě OPE-1 s jedním záznamem bylo použito Dublinské jádro. Tuto skutečnost bylo možné zjistit díky přítomnosti výrazu „rft_val_fmt=info%3Aofi%2Ffmt%3Akev%3Amtx%3Adc“. Zápis odpovídá výrazu „info:ofi/fmt:kev:mtx:dc“, místo znaků „:“ a „/“ jsou však použita vyjádření formou procent, konkrétně výrazy „%3A“ a „%@F“, které jsou pro přehlednost v úvodním výrazu barevně zvýrazněny. Uvozující část před rovnítkem oddělujícím typ údaje od jeho hodnoty, tj. „rft_val_fmt“, udává, že se jedná o formát hodnoty referentu, tj. zdroje odkazovaného na síti.

V záznamu je použito celkem devět různých prvků, a to „rft.title“, „rft.creator“, „rft.description“, „rft.date“, „rft.format“, „rft.identifier“, „rft.language“, „rft.subject“ a „rft.type“. Poslední dva jmenované prvky se v záznamu vyskytují vždy dvakrát, ostatní prvky pouze jednou.

Pokud se týká autorských údajů, byl v sadě k dispozici pouze jeden výskyt, viz tab. 32

Tab. 32: Pole s autorskými údaji ve formátu OpenURL COinS

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
OPE-1	&rft.creator=	1	1 výskyt jména v invertované podobě (znaky „%2C“ nahrazují čárku, místo mezery je použit znak „+“)	Elaieess %2C+Ramadan

Úplný ani zkrácený název časopisu se v analyzovaném záznamu nevyskytoval.

Poznámky

Existuje verze formátu OpenURL v jazyce XML [188, s. 47-72]. Konkrétně pro formát OpenURL COinS však bylo záměrně zvoleno využití prvku „span“ a nikoliv např. využití XML se jmennými prostory, které by bylo zahrnuto do formátu XHTML, neboť byly očekávány možné problémy vyplývající z případné

nemožnosti přistupovat k datům prostřednictvím Javascriptu. Rovněž bylo přihlíženo k požadavku, aby formát byl v souladu s jazykem HTML, nikoliv XHTML.

4.5 Formáty využívající jazyk XML

4.5.1 MARCXML

Specifikace

Specifikace formátu MARCXML je k dispozici v rámci oficiální webové prezentace formátu, jíž spravuje Kongresová knihovna ve Washingtonu. Kromě vlastního XML schématu jsou k dispozici např. také příklady použití formátu a nástroje pro práci s formátem [196], [197].

Formát vychází z formátu MARC 21, jehož specifikace je rovněž dostupná na webu Kongresové knihovny, a zahrnuje všechny části tohoto formátu (nejenom MARC 21 pro bibliografická data).

Údaje o aktualizaci formátu jsou k dispozici v rámci schématu v prvku „documentation“ – verze 1.0 formátu byla aktuální k 21. květnu 2002, verze 1.1 ke 4. srpnu 2003, nejnovější verzí je verze 1.2 z 21. května 2009.

Syntax

Syntax z velké části určuje skutečnost, že MARCXML je XML formát. Způsob práce s bílými znaky, význam konce řádku a možnost využití řádkování v delších textech jsou dány specifikací jazyka XML. V případě práce s bílými znaky můžeme upřesnit, že ve formátu MARCXML je v některých případech (zejména se jedná o hodnoty prvků „leader“ a „controlfield“ a také prvku „subfield“) tyto znaky třeba zachovat. Pro tento účel je ve schématu využit prvek „xsd:whiteSpace“ a hodnota „preserve“.

Specifikace XML také upřednostňuje konkrétní kódování, a to UTF-8. V případě datové sady MAR-1 bylo zjištěno právě toto kódování (viz tab. 33).

Tab. 33: Kódování souboru ve formátu MARCXML

ID sady	Kódování
MAR-1	UTF-8

Soubory ve formátu MARCXML začínají XML deklarací udávající verzi XML a použité kódování. Vlastní data jsou pak zahrnuta v prvku „collection“, jehož součástí jsou prvky „record“ a jim dále podřazené prvky.

V rámci jednotlivých záznamů je možné opakovat prvky s výjimkou prvku „leader“. Ten by se měl v záznamu vyskytovat pouze jednou.

Sémantika

Označení jednotlivých typů údajů vychází z terminologie formátů typu MARC. Ve formátu MARCXML jsou využívány prvky „collection“, „record“, „leader“, „controlfield“, „datafield“ a „subfield“, tj. celkem šest prvků. U prvků „leader“ a „controlfield“ nejsou používány atributy, u ostatních zmíněných prvků naopak ano. V případě prvků „collection“ a „record“ je použití atributů volitelné, u prvků „datafield“ a „subfield“ se setkáváme s povinnými i volitelnými atributy. Typicky jsou u prvku „datafield“ používány atributy „tag“, „ind1“ a „ind2“ a u prvku „subfield“ pak atribut „code“.

Pokud se jedná o otázku rozlišování editorů od autorů a zkrácených názvů časopisů od úplných, lze konstatovat, že je možné.

Autory a editory lze odlišit prostřednictvím kódů rolí, které je možné uvést v podpoli 4 v polích, v nichž jsou uváděny autorské údaje v selekční podobě. Zejména se jedná o pole 100 a 110 (primární autorská odpovědnost personální a korporativní) a o pole 700 a 710 (sekundární autorská odpovědnost personální a korporativní). Podpole 4 může v praxi nabývat mj. hodnot „aut“ a „edt“ [144].

Pro zkrácený název je speciálně určeno pole 210, které prostřednictvím druhého indikátoru umožňuje upřesnit, zda se jedná o zkrácený klíčový název podle ISSN (hodnota „#“) nebo o název zkrácený podle jiných pravidel (hodnota „0“). Ve druhém případě je pak možné do podpole 2 doplnit kód zdroje, který byl využit pro zkrácování [198]. K zachycení názvů jsou určena i další pole (např. pole 222 pro klíčový název). Při tvorbě záznamů časopiseckých článků je však pro zachycení údajů o zdrojovém dokumentu využíváno pole 773, i to (vedle podrobných názvových údajů) umožňuje zachycení zkráceného názvu (podpole „p“) či např. identifikátoru CODEN (podpole „y“).

Pravidla pro tvorbu obsahu

Pokud se týká pravidel pro tvorbu obsahu, je formát MARC 21 (a tedy i formát MARCXML) v praxi používán především spolu s pravidly AACR2. Právě ta potom udávají způsob vyplňování údajů. Tato pravidla jsou velmi podrobná, ne vždy však zcela jednoznačná.

V datové sadě MAR-1 bylo pro zachycení názvu zdrojového dokumentu použito pole „datafield“ s atributem „tag“ s hodnotou „773“ (je ovšem uveden úplný, ne zkrácený název):

```
<datafield tag="773" ind1=" " ind2=" " >
  <subfield code="a">10.1007/s11786-010-0029-2</subfield>
  <subfield code="c">20</subfield>
  <subfield code="o">Mathematics in Computer Science 3, 3 (2010)                20</
subfield>
  <subfield code="p">Mathematics in Computer Science</subfield>
  <subfield code="v">3</subfield>
  <subfield code="y">2010</subfield>
</datafield>
```

S autorskými údaji se v datové sadě MAR-1 setkáváme v poli „datafield“ s atributem „tag“ s hodnotou „100“:

```
<datafield tag="100" ind1=" " ind2=" " >
  <subfield code="a">Bouche, Thierry</subfield>
  <subfield code="u">IF, CCDNM</subfield>
</datafield>
```

4.5.2 MODS

Specifikace

Specifikace formátu MODS je k dispozici v rámci oficiální webové prezentace formátu, kterou rovněž spravuje Kongresová knihovna ve Washingtonu. Ta navíc kromě samotné specifikace formátu v XML schématu nabízí řadu dalších užitečných nástrojů a dokumentů, např. mapovací tabulky prvků formátu do dalších formátů, styly v jazyce XSLT¹¹⁷ určené pro konverzi mezi různými formáty z nebo do

¹¹⁷ XSLT je zkratka výrazu *Extensible Stylesheet Language Transformations* (rozšiřitelný stylový jazyk).

formátu MODS, návod pro tvorbu záznamů ve formátu MODS či registr implementací [199].

Údaje o aktualizaci formátu jsou velmi dobře dostupné, neboť jednotlivé verze formátu jsou číslovány – k dubnu 2010 byla aktuální verze formátu (resp. XML schématu) 3.3 – právě té je věnována pozornost dále v textu. (Součástí XML schématu jsou navíc podrobné komentáře, které vysvětlují změny oproti předcházející verzi formátu).

Syntax

Syntax je do značné míry dána skutečností, že se jedná o XML formát. Ve schématu je předepsáno kodování UTF-8 (toto kódování bylo rozpoznáno i v datové sadě MOD-1, viz tab. 34). Na začátku souboru se pochopitelně nachází XML deklarace, po ní následuje v případě sady záznamů kořenový prvek „modsCollection“, v případě jednotlivého záznamu kořenový prvek „mods“. Záznam uzavírají koncové tagy prvků „modsCollection“, resp. „mods“.

Tab. 34: Kódování souboru ve formátu MODS

ID sady	Kódování
MOD-1	UTF-8

Způsob práce s bílými znaky, význam konce řádku a možnost řádkování jsou přesně dány specifikací jazyka XML.

Pokud se týká řešení případů více hodnot jednoho typu údaje, je ve formátu MODS možné opakovat všechny prvky na nejvyšší úrovni (s výjimkou prvku „recordInfo“).

Sémantika

Jednotlivé typy údajů jsou označovány výrazy vycházejícími z původních jednoslovných či víceslovných výrazů v angličtině. V případě víceslovných vyjádření jsou první písmena druhých, popř. dalších slov vždy velká, aby bylo možné snadno zjistit název prvku.

Formát obsahuje dva kořenové prvky („mods“ a „modsCollection“). Dále je na nejvyšší úrovni jeho součástí celkem dvacet prvků:

titleInfo	note
name	subject
typeOfResource	classification
genre	relatedItem
originInfo	identifier
language	location
physicalDescription	accessCondition
abstract	part
tableOfContents	extension
targetAudience	recordInfo

Přehled dalších prvků, které jsou jim podřazeny, podává tab. 35. V případě, že se u daného prvku na nejvyšší úrovni setkáváme nejenom s prvky přímo podřazenými tomuto prvku, ale také s třetí, resp. čtvrtou hierarchickou úrovní, jsou tyto prvky uvedeny v závorce u podřazeného prvku, resp. v závorce po středníku. Do tabulky jsou také zařazeny pouze názvy atributů prvků na nejvyšší úrovni (přehled všech atributů je k dispozici v rámci XML schématu nebo přehledu prvků a atributů formátu dostupného ve formátu HTML).

Tab. 35: Přehled prvků, jejich atributů a podřazených prvků ve formátu MODS

Název prvku	Názvy atributů	Názvy podřazených prvků
titleInfo	ID, xlink, lang, xml:lang, script, transliteration, type, authority, displayLabel	title, subTitle, partNumber, partName, nonSort
name	ID, xlink, lang, xml:lang, script, transliteration, type, authority	namePart, displayForm, affiliation, role (roleTerm), description
typeOfResource	collection, manuscript	–
genre	lang, xml:lang, script, transliteration, authority, type	–
originInfo	lang, xml:lang, script, transliteration	place (placeTerm), publisher, dateIssued, dateCreated, dateCaptured, dateValid, dateModified, copyrightDate, dateOther, edition, issuance, frequency

Název prvku	Názvy atributů	Názvy podřazených prvků
language	objectPart	languageTerm
physicalDescription	lang, xml:lang, script, transliteration	form, reformattingQuality, internetMediaType, extent, digitalOrigin, note
abstract	xlink, lang, xml:lang, script, transliteration, displayLabel, type	–
tableOfContents	xlink, lang, xml:lang, script, transliteration, displayLabel, type	–
targetAudience	lang, xml:lang, script, transliteration, authority	–
note	ID, xlink, lang, xml:lang, script, transliteration, displayLabel, type	–
subject	ID, xlink, lang, xml:lang, script, transliteration, authority	topic, geographic, temporal, titleInfo, name, geographicCode, genre, hierarchicalGeographic (continent, country, province, region, state, territory, county, city, island, area, extraterrestrialArea, citySection), cartographics (scale, projection, coordinates), occupation
classificaton	lang, xml:lang, script, transliteration, authority, edition, displayLabel	–
relatedItem	ID, xlink, displayLabel, type	jakýkoliv prvek formátu MODS podle schématu
identifier	lang, xml:lang, script, transliteration, type, displayLabel, invalid	–
location	–	physicalLocation, shelfLocator, url, holdingSimple (copy information; form, sublocation, shelfLocator, electronicLocator, note, enumerationAndChronology), holdingExternal
accessCondition	xlink, lang, xml:lang, script, transliteration, displayLabel, type	–
part	ID, type, order	detail (number, caption, title), extent (start, end, total, list), date, text

Název prvku	Názvy atributů	Názvy podřazených prvků
extension	–	–
recordInfo	lang, xml:lang, script, transliteration	recordContentSource, recordCreationDate, recordChangeDate, recordIdentifier, recordOrigin, languageOfCataloging (languageTerm), descriptionStandard

Celkový počet prvků na druhé hierarchické úrovni je 55. Započítáme-li prvky, které je možné podřadit prvku „relatedItem“ (tj. dvacet prvků na nejvyšší úrovni), získáme počet 75 prvků. Na třetí úrovni pracuje formát s 26 prvky a na čtvrté úrovni se šesti prvky.

Na druhé hierarchické úrovni formát pracuje s celkem patnácti různými názvy atributů, přičemž řadu z nich lze použít u více než jednoho prvku.

Pokud se týká rozlišování autorů a editorů, je možné konstatovat, že formát MODS tuto oblast má pokrytou velmi podrobně – u prvku „name“ lze totiž prostřednictvím podřazeného prvku „role“, resp. jemu podřazeného prvku „roleTerm“ a atributu „authority“ určit roli prostřednictvím relátorů a kódů formátu MARC 21. Seznam těchto relátorů a kódů je velmi obsáhlý, kromě rolí „Author“ (kód „aut“) či „Editor“ (kód „edt“) nabízí ještě 211 dalších rolí a jejich kódů.

V případě názvů časopisů vyjdeme z příkladu záznamu časopiseckého článku dostupného na oficiální webové prezentaci formátu [200]. V tomto případě je jako hodnota prvku „title“ (tento prvek je podřazen prvku „titleInfo“, jehož nadřazeným prvkem je prvek „relatedItem“) uveden pouze úplný název časopisu (hodnota „International Journal of Urban and Regional Research“). Z XML schématu je však zřejmé, že formát umožňuje i ukládání zkrácených názvů časopisů – prvek „titleInfo“ totiž může mít atribut „type“ s hodnotou „abbreviated“ (byť návod pro implementaci záznamů ve formátu MODS určených pro sdílení mezi různými systémy tento případ explicitně nezmiňuje [201], nečiní tak ani odkazovaný dokument [202]).

Pravidla pro tvorbu obsahu

Formát MODS umožňuje uvádění dat ve velmi granulované podobě, proto je místo speciálních pravidel pro tvorbu obsahu možné vycházet z XML schématu

formátu, dále také z návodů a příkladů dostupných v rámci webové prezentace formátu [203].

Pro formát MODS je typické využití řady slovníků. Např. v případě atributu „authority“ u prvku „name“ je explicitně uveden odkaz na stránku *Name and Title Authority Source Codes* [204]. V tomto případě a případech analogických je tedy přesně určeno, jakým způsobem vyplňovat hodnoty prvků, resp. umožňuje jednoznačné uvedení zdroje dat.

Typická je rovněž návaznost na další existující standardy, např. v případě atributu „encoding“, který je užit u prvků týkajících se dat a časů na druhé hierarchické úrovni, jsou jako možné hodnoty uvedeny „w3cdtf“, „iso8601“ a „marc“.¹¹⁸

Datová sada MOD-1 (konkrétně ve verzi 3.0 formátu) obsahovala pouze jeden záznam, název časopisu (zkrácený ani úplný) nebyl v záznamu uveden. Autorské údaje naopak tvořily součást záznamu, konkrétně se jednalo o první jméno a příjmení autora. Každý z údajů se vyskytoval v samostatném prvku, upřesnění obsahu bylo realizováno prostřednictvím hodnot atributů:

```
<mods:name type="personal">
  <mods:namePart type="given">Ramadan</mods:namePart>
  <mods:namePart type="family">Elaiess</mods:namePart>
  <mods:role>
    <mods:roleTerm type="text">author</mods:roleTerm>
  </mods:role>
</mods:name>
```

4.5.3 NLM

Specifikace

Specifikace XML formátu NLM je k dispozici v podobě DTD dostupných na webu NLM [206]. Kromě vlastních DTD je dostupná i doprovodná dokumentace [207].

¹¹⁸ V této souvislosti je vhodné doplnit, že Kongresová knihovna ve Washingtonu se v současné době snaží o vytvoření a následné prosazení (např. formou standardu) specifikace rozšířeného formátu data a času (*Extended Date/Time Format*, EDTF) [205].

Údaje o aktualizaci formátu jsou k dispozici jednak v DTD formou komentáře, jednak v rámci doprovodné dokumentace. Obecně lze konstatovat, že formát je každoročně aktualizován, přičemž je reflektován produkční rok databáze MEDLINE, resp. celého systému PubMed. Aktuální verze je tedy pro rok 2010. V dokumentaci je pak uváděno datum poslední revize a aktualizace konkrétní webové stránky, dostupné jsou také údaje o míře trvalosti dokumentu. Průběžně jsou prováděny dílčí změny, rovněž ty jsou podrobně a přehledně dokumentovány.

NLM provozuje řadu informačních systémů a spolupracuje s externími subjekty. Proto je na tomto místě potřeba zdůraznit, že se v praxi je možné setkat s různými variantami formátu, resp. s více formáty dostupnými pod hlavičkou NLM. V sadě NLM-1 byl použit formát „PubMedArticle“, v sadě NLM-2 pak formát označovaný jako „Journal Publishing Tag Set“ (Sada tagů pro publikování časopisů) [208]. Formát „PubMedArticle“ je velmi podobný popisovanému formátu (používá stejné prvky), formát „Journal Publishing Tag Set“ je konstruován odlišně. V dalším textu je proto věnována pozornost zejména sadě NLM-1, sada NLM-2 je uvedena pouze doplňkově.

Syntax

Jedná se o XML formát, proto je za implicitní kódování možné považovat kódování UTF-8. Toto kódování bylo identifikováno i v případě dvou datových sad NLM-1 a NLM-2. V sadě NLM-1 nebylo určeno v deklaraci, k rozpoznání kódování byl využit textový editor PSPad (viz tab. 36).

Tab. 36: Kódování souboru ve formátu NLM

ID sady	Kódování
NLM-1	UTF-8
NLM-2	UTF-8

V záznamech ve formátu NLM je po XML deklaraci uváděn prvek „MedlineCitationSet“, resp. v případě exportu dat ze systému PubMed prvek „PubMedArticleSet“.

Způsob práce s bílými znaky, význam konce řádku a možnost řádkování jsou dány specifikací jazyka XML.

Řadu prvků je možné v případě potřeby opakovat, např. prvek „Author“ podřazený prvku „AuthorList“ se může vyskytovat vícekrát.

Sémantika

Formát obsahuje celkem 85 prvků s názvy vycházejícími z přirozeného jazyka. Na nejvyšší úrovni nachází prvek „MedlineCitationSet“, jemu je podřazen prvek „MedlineCitation“. Na hierarchicky nižší úrovni se pak nalézají 19 prvků:

PMID	MeshHeadingList
DateCreated	NumberOfReferences
DateCompleted	PersonalNameSubjectList
DateRevised	OtherID
Article	OtherAbstract
MedlineJournalInfo	KeywordList
ChemicalList	SpaceFlightMission
CitationSubset	InvestigatorList
CommentsCorrectionsList	GeneralNote
GeneSymbolList	

Z našeho pohledu hraje nejpodstatnější roli prvek „Article“ a prvky jemu dále podřazené. Jejich přehled podává tab. 37 (prvky „StartPage“ a „EndPage“ jsou v DTD uvedeny, ale jsou rezervovány pro budoucí užití).

Tab. 37: Přehled prvků podřazených ve formátu NLM prvku „Article“, jejich atributů a podřazených prvků

Název prvku	Názvy atributů	Názvy podřazených prvků
Journal	–	ISSN, JournalIssue, Title, ISOAbbreviation
ArticleTitle	–	–
Pagination	–	StartPage, EndPage, MedlinePgn
ELocationID	EldType	
Abstract	–	AbstractText, CopyrightInformation
Affiliation	–	–
AuthorList	ValidYN	Author
Language	–	–
DataBankList	CompleteYN	DataBank
GrantList	–	Grant

Název prvku	Názvy atributů	Názvy podřazených prvků
PublicationTypeList	–	PublicationType
Vernacular Title	–	–
ArticleDate	DateType	Year, Month, Day

Ve formátu je tedy k dispozici prvek „Author“. Jemu mohou být ještě podřazené prvky „LastName“, „ForeName“, „Initials“, „Suffix“ a v budoucnu se bude jednat i o prvek „NameID“. Pro korporativní autorství navíc lze použít prvek „CollectiveName“. Zvláštní prvek pro vyjádření role editora není k dispozici.

Odlišení různých názvů časopisů je možné realizovat prostřednictvím prvků „Title“ (úplný název) a „ISOAbbreviation“ (zkrácený název podle normy ISO) podřazených prvku „Journal“. Pro zkrácený název podle databáze MEDLINE je určen prvek „MedlineTA“.

Pravidla pro tvorbu obsahu

Pravidla pro tvorbu obsahu jsou naznačena v dokumentaci, která je relativně podrobná. To přispívá k relativně vysoké míře jednoznačnosti.

Pokud se týká názvů zdrojových dokumentů, je v datové sadě NLM-1 díky přítomnosti prvků „Title“, „ISOAbbreviation“ a „MedlineTA“ odlišen typ názvu, resp. jeho zkrácení (viz tab. 38).

Tab. 38: Prvky s názvy zdrojových dokumentů ve formátu NLM

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
NLM-1	Title	25	25 výskytů úplných názvů zdrojových dokumentů, v některých případech doplněných např. o místo vydání; názvy zdrojových dokumentů s výjimkou prvního písmene názvu a případných vlastních jmen v názvu začínají malými písmeny	European radiology Neoplasia (New York, N.Y.) Forensic science international Journal of chromatography. A
NLM-1	ISOAbbreviation	21	21 výskytů zkrácených názvů zdrojových dokumentů (krácení provedeno v souladu s příslušnou normou)	Neoplasia Forensic Sci. Int. J Chromatogr A

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
			ISO)	
NLM-1	MedlineTA	25	25 výskytů zkrácených názvů zdrojových dokumentů (krácení provedeno podle způsobu používaného v databázi MEDLINE)	Eur Radiol Neoplasia Forensic Sci Int J Chromatogr A
NLM-2	journal-title	1	1 výskyt úplného názvu zdrojového dokumentu	Mathematics in Computer Science
NLM-2	abbrev-journal-title	1	1 výskyt úplného názvu zdrojového dokumentu	Mathematics in Computer Science

Podobně v případě prvků s autorskými údaji jsou díky přítomnosti dílčích prvků dobře odlišeny jednotlivé části jmen autorů (viz tab. 39).

Tab. 39: Prvky s autorskými údaji ve formátu NLM

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
NLM-1	LastName	105	105 výskytů příjmení (v případech, kdy je součástí příjmení prefix typu „van“ či „von der“, jsou prefixy uvedeny v rámci příjmení, tj. v přímém pořadí)	Rodríguez Gonzáles van Hemert Bashkirov von der Lieth
NLM-1	ForeName	105	9 výskytů iniciál prvního jména, popř. dalších jmen 96 výskytů úplného prvního jména, popř. středního jména či jmen nebo kombinace alespoň jedné úplné podoby jména doplněné zkráceným tvarem (většinou středního jména; v případě zkrácených tvarů jsou jména uváděna bez teček a s mezerami)	Didier Vladimir I E Aubrey I G L M M D B S
NLM-1	Initials	105	105 výskytů iniciály prvního jména nebo iniciály prvního jména a středního jména či jmen (mezi iniciálami nejsou mezery)	P HY ED JM MD
NLM-1	Suffix	1	1 výskyt sufixu v obvykle používané zkrácené podobě	3rd
NLM-1	Collective Name	1	1 výskyt korporativního autora (pracovní skupiny)	Evidence-Based Urology Working Group
NLM-2	surname	1	1 výskyt příjmení	Bouche
NLM-2	given-	1	1 výskyt úplné podoby prvního jména	Thierry

ID sady	Název pole	Počet výskytů v sadě	Komentář k obsahu pole	Příklady
	names			

4.6 Dílčí závěry

Pokud se týká citačních formátů, lze konstatovat, že různé informační systémy nabízejí stažení dat v různých formátech. Často se vyskytují např. formáty RIS, EndNote, RefWorks či BibTeX. V prvních třech případech se jedná o formáty určené primárně jako formáty vstupních dat do citačních manažerů Reference Manager, EndNote a RefWorks, ve čtvrtém případě se jedná o formát, který byl původně vyvinut pro potřeby typografického systému LaTeX. Objevují se také formáty v jazyce XML, např. MARCXML, MODS či NLM, které jsou používány mj. pro uložení a přenos citačních dat. Průzkum výskytu jednotlivých citačních formátů v celkem čtrnácti informačních systémech (Web of Science, Scopus, Nature, Science, SpringerLink, ScienceDirect, Wiley InterScience, PubMed, EBSCOhost, ProQuest Central, WorldCat, bibliotek.dk, CERN Document Server a E-LIS) ukázal, že výstupy v citačních formátech v jazyce XML jsou však zatím spíše výjimkou než pravidlem.

Přesto lze konstatovat, že jazyk XML se postupně prosazuje jako jazyk strojově srozumitelných citačních formátů. Toto tvrzení významně podporují zvláště následující dvě skutečnosti:

- 1) řada citačních formátů, které vznikly v době před vznikem a rozšířením jazyka XML, byla doplněna o svou verzi právě v tomto jazyce,
- 2) nově vznikají citační formáty v jazyce XML, které nejsou pouze jednoduchou obdobou dřívějších formátů nezaložených na XML, ale výrazněji využívají možností XML.

Různé informační systémy používají citační formáty různým způsobem, navíc u všech formátů není veřejně dostupná jejich podrobná specifikace (jako příklad může posloužit formát ISI). Praxe také ukazuje, že je-li ve specifikaci uvedeno kódování, které nevyhovuje současným potřebám, není specifikace dodržována. Příkladem může být rozšířená znaková sada IBM jako předepsané kódování formátu RIS. Ta je však často nahrazována kódováním UTF-8.

Značně problematický je také nejednotný způsob uvádění vlastních citačních dat, kdy jsou například jména autorů uváděna někdy v invertované a někdy v přímé podobě. Podobně ne vždy jsou rozlišovány úplné názvy časopisů od názvů zkrácených.

Neuplatňování jednotných standardů komplikuje možnost automatického zpracování citačních dat. S daty je mnohdy možné pracovat pouze po provedení parsování. Jeho realizace není jednoduchá, navíc nepřináší vždy dostatečně přesné výsledky, může tedy docházet ke zkreslení původních dat.

5. Citační manažery

5.1 Úvod

V souvislosti s rozvojem informačních a komunikačních technologií se významně rozvíjejí také tzv. citační manažery, tedy softwarové nástroje, které umožňují do určité míry automatizovat proces citování při psaní odborného či vědeckého textu. Pomáhají při řešení situace, kdy autor publikuje v časopisech, které používají rozdílné citační styly. Prostřednictvím citačních manažerů lze ukládat data o citovaných dokumentech, často též vytvářet automaticky seznamy použitých zdrojů ve zvoleném citačním stylu a rovněž vkládat citace do konkrétních míst v textu, popř. využívat další dostupné funkce.

Existuje řada různých citačních manažerů, ať již jednoduchých nebo komplexních, dostupných jako desktopová aplikace nebo jako online služba či nástroj využitelný v rámci internetového prohlížeče. Zatímco ještě před několika lety se typicky jednalo o desktopovou aplikaci, nyní se těžiště přesouvá právě směrem k online službám či doplňkům do internetových prohlížečů, které se z hlediska uživatele stávají integrální součástí jeho činností realizovaných ve webovém prostředí.

Mezi citační manažery, s nimiž lze pracovat prostřednictvím internetových prohlížečů, patří mj. Zotero (doplňek do Firefoxu) [99] nebo Connotea [209]. Tyto dva konkrétní příklady citačních manažerů zároveň ukazují další trend – vedle tradičních komerčních citačních manažerů (např. ProCite či EndNote) se objevují také nástroje s otevřeným zdrojovým kódem, které je možné je využívat zdarma (tj. nejsou vázány na platbu licenčních poplatků) a navíc se v případě zájmu lze aktivně podílet na jejich dalším rozvoji. Podrobnější informace o citačních manažerech jsou k dispozici např. v [4].

Pro import a export dat z citačních manažerů jsou většinou k dispozici různé citační formáty (RIS, BibTeX apod.), při přenosu dat z jednoho citačního manažeru do druhého však může dojít k nepřesnostem při interpretaci dat.

V následujícím textu jsou podrobněji představeny čtyři vybrané citační manažery. Jedná se o dva nástroje s otevřeným zdrojovým kódem (Zotero

a Connotea) a o dva nástroje bez otevřeného zdrojového kódu, s nimiž je však možné pracovat ve webovém prostředí (Citace 2.0 a CiteULike).

Po uvedení základních údajů o daném softwaru (doba vzniku, nejnovější verze, popř. další doplňující údaje) jsou zachyceny odpovědi na následující otázky (bylo-li je v daném případě možné získat):

- Ve kterých formátech je možné do citačního manažeru importovat data?
- Ve kterých formátech je naopak možné data exportovat?
- Které z výše uvedených formátů jsou v jazyce XML?
- Jakou podobu má interní formát uložení dat v daném manažeru?
- Jakými způsoby je možné do citačního manažeru automaticky přebírat data?

5.2 Zotero

Zotero je citační manažer, který byl vytvořen Centrem pro historii a nová média na Univerzitě George Masona (*Center for History and New Media, George Mason University*), a to za finanční podpory Ústavu pro muzejní a knihovní služby Spojených států (*United States Institute of Museum and Library Services*), Nadace Andrewa W. Mellona (*Andrew W. Mellon Foundation*) a Nadace Alfreda P. Sloana (*Alfred P. Sloan Foundation*) [99], [210].

Verze 1.0 softwaru Zotero byla zpřístupněna v roce 2006, ke květnu 2010 je aktuální verze 2.0.3.

Software Zotero má podobu zásuvného modulu, který je možné instalovat v prohlížeči Mozilla Firefox.

Formátů, v nichž je možné do Zotera importovat data, je celkem šest – jedná se o formáty Zotero RDF, MODS, BibTeX, RIS, Refer/BibIX a o formát Unqualified Dublin Core RDF. Exportní formáty zahrnují jednak šest již zmíněných importních formátů, jednak formát Wikipedia Citation Template.

V jazyce XML jsou celkem tři z těchto formátů (Zotero RDF, MODS a Unqualified Dublin Core RDF). Ukázky jednoho záznamu exportovaného v těchto třech XML formátech jsou k dispozici na obr. 22, 23 a 24.

```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:bib="http://purl.org/net/biblio#" xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:z="http://www.zotero.org/namespaces/export#">
  <bib:Manuscript rdf:about="http://eprints.rclis.org/7252/">
    <dc:date>2005</dc:date>
    <dc:identifier>
      <dcterms:URI>
        <rdf:value>http://eprints.rclis.org/7252/</rdf:value>
      </dcterms:URI>
    </dc:identifier>
    <dc:subject>DINI certificate</dc:subject>
    <dc:subject>Deutsche Initiative für Netzwerk Information</dc:subject>
    <dc:subject>German universities</dc:subject>
    <dc:subject>Institutional repositories</dc:subject>
    <dc:subject>Nestor</dc:subject>
    <dc:subject>Network of Expertise in Long-Term Storage of Digital Resources</dc:subject>
    <dc:subject>Open access in Germany</dc:subject>
    <dc:subject>archiving of scientific documents</dc:subject>
    <dc:subject>certification</dc:subject>
    <dc:subject>electronic publishing</dc:subject>
    <dc:subject>interoperability</dc:subject>
    <dc:title>Certification and beyond : DINI open access activities in Germany</dc:title>
    <dcterms:abstract>Local publication servers are common and at the same time highly
      fragmented in Germany. To bring them to greater success it is necessary to standardize
  
```

Obr. 22: Část záznamu e-printu ve formátu Zotero RDF

```

<?xml version="1.0"?>
<modsCollection
  xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd"
  xmlns="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <mods>
    <titleInfo>
      <title>Certification and beyond : DINI open access activities in Germany</title>
    </titleInfo>
    <typeOfResource>text</typeOfResource>
    <genre authority="local">manuscript</genre>
    <name type="personal">
      <namePart type="family">Dobratz</namePart>
      <namePart type="given">Susanne</namePart>
      <role>
        <roleTerm type="code" authority="marcrelator">aut</roleTerm>
      </role>
    </name>
    <name type="personal">
      <namePart type="family">Scholze</namePart>
      <namePart type="given">Frank</namePart>
      <role>
        <roleTerm type="code" authority="marcrelator">aut</roleTerm>
      </role>
    </name>
    <originInfo>
      <dateCreated>2005</dateCreated>
    </originInfo>
    <location>
      <url>http://eprints.rclis.org/7252/</url>
    </location>
    <abstract>Local publication servers are common and at the same time highly fragmented in
      Germany. To bring them to greater success it is necessary to standardize further developments.
  
```

Obr. 23: Část záznamu e-printu ve formátu MODS

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://eprints.rclis.org/7252/">
    <dc:creator>Dobratz, Susanne</dc:creator>
    <dc:creator>Scholze, Frank</dc:creator>
    <dc:date>2005</dc:date>
    <dc:title>Certification and beyond : DINI open access activities in Germany</dc:title>
    <dc:type>manuscript</dc:type>
  </rdf:Description>
</rdf:RDF>

```

Obr. 24: Část záznamu e-rintu ve formátu Unqualified Dublin Core RDF

Data uložená uživatelem v citačním manažeru Zotero jsou umístěna v adresáři obsahujícím profil uživatele Firefoxu [211]. Většina dat (s výjimkou úplných textů) je uložena v souboru s názvem „zotero.sqlite“¹¹⁹ v podobě, která je pro uživatele-člověka pouze obtížně srozumitelná (viz obr. 25).

	0001	0203	0405	0607	0809	0A0B	0C0D	0E0F	0123456789ABCDEF
000000	5351	4C69	7465	2066	6F72	6D61	7420	3300	SQLite format 3.
000010	0400	0101	0040	2020	0000	009F	0000	0000@ ...ž....
000020	0000	0000	0000	0000	0000	01A6	0000	0001!....
000030	FFFF	E0C0	0000	007E	0000	0001	0000	0000	' 'řř....~.....
000040	0000	0000	0000	0000	0000	0000	0000	0000
000050	0000	0000	0000	0000	0000	0000	0000	0000
000060	0000	0000	0503	B000	7001	6802	0000	008C° .p.h....š
000070	03FB	03F6	03F1	03EC	03E7	0253	03E2	03D8	.ú.ö.ń.ě.ç.s.â.ř
000080	03D3	03CE	03C9	03C4	03BF	03BA	0199	03AB	.ó.î.é.ä.ž.š.™.«
000090	03A6	03A1	039C	0396	0390	038A	0384	037E	.! . ˇ . ś . - . . . š . „ . ~
0000A0	0378	0372	036C	0366	0360	035A	0354	034E	.x.r.l.f.`.Z.T.N
0000B0	0348	0342	033C	0336	0330	032A	0324	031E	.H.B.<.6.0.*.š..
0000C0	0318	0312	030C	0306	0300	02FA	02F4	02EEú.ô.î
0000D0	02E8	02E2	02DC	02D6	02D0	02CA	02C4	02BE	.č.â.Û.Ö.Đ.Ě.Ä.I
0000E0	02B8	02B2	02AC	02A6	02A0	029A	0294	028E š . " . ž

Obr. 25: Úvodní část souboru s daty uloženými citačním manažerem Zotero

Zotero využívá ve své databázi principů objektově orientovaného programování, které vychází ze způsobu programování v jazyce Java. Interně jsou data uložena v databázi SQL¹²⁰, přičemž data jsou umístěna ve velkém počtu vzájemně provázaných tabulek [12] (viz obr. 26).

¹¹⁹ SQLite je běžně využívaná relační databáze s otevřeným zdrojovým kódem. Je často využívána autory softwaru, kteří potřebují databázi, u níž neočekávají příliš velkou velikost uložených dat.

¹²⁰ SQL je zkratka výrazu *Structured Query Language* (strukturovaný dotazovací jazyk).

```

CREATE TABLE itemTypes (
  itemTypeID INTEGER PRIMARY KEY,
  typeName TEXT,
  templateItemTypeID INT,
  display INT DEFAULT 1 -- 0 == hide, 1 == display, 2 == primary
);

-- Populated at startup from itemTypes and customItemTypes
DROP TABLE IF EXISTS itemTypesCombined;
CREATE TABLE itemTypesCombined (
  itemTypeID INT NOT NULL,
  typeName TEXT NOT NULL,
  display INT DEFAULT 1 NOT NULL,
  custom INT NOT NULL,
  PRIMARY KEY (itemTypeID)
);

```

Obr. 26: Část dotazu v jazyce SQL pro vytvoření tabulky „itemTypes“

Pokud se týká automatického přebírání dat do Zotera, lze konstatovat, že je na velmi vyspělé úrovni. Data je možné automaticky získávat prostřednictvím ikony knihy, která se objevuje v pravé části adresového řádku prohlížeče Mozilla Firefox, pokud si uživatel zobrazuje příslušná data na webech, z nichž je Zotero umí automaticky přebírat prostřednictvím tzv. překladačů (*translators*) [212]. Data je také možné získávat prostřednictvím identifikátorů (ISBN, DOI či PMID) – kvalita dat závisí vždy na jejich zdroji (na základě ISBN jsou data nahrávána z katalogu WorldCat, k získání dat o zdrojích reprezentovaných identifikátorem DOI je použit systém CrossRef). Za automatizovaný způsob získání dat lze pokládat také možnost importovat do Zotera data z jiných softwarů (v některém z výše zmíněných formátů pro import) [213]. Obdobou tohoto způsobu získání dat je také stažení citačních dat ve vhodném formátu např. z některé digitální knihovny (samozřejmě pokud by tato data nebylo možné získat nejjednodušším možným způsobem, tj. prostřednictvím ikony knihy v adresovém řádku). Software Zotero navíc umožňuje archivaci webových stránek či přiřazování dalších souborů k záznamům jednotlivých zdrojů.

5.3 Connotea

Citační manažer a nástroj pro vytváření sdílených záložek Connotea [209] je produktem Nature Publishing Group. Jedná se o online nástroj, který byl spuštěn na konci roku 2004 [214]. Na rozdíl od softwaru Zotero není k využívání Connotey potřeba instalovat dodatečné zásuvné moduly do prohlížeče, je pouze potřeba si

vytvořit uživatelský účet (jeho vytvoření, resp. registrace na serveru Connotey je zdarma). Aktuální stabilní verze softwaru je označena jako verze 1.8 [13].

Connotea nabízí celkem osm formátů dat pro import – jsou jimi formáty Firefox Bookmarks, RIS, Endnote (Refer), Endnote (XML) [Experimental], BibTeX, ISI Web of Knowledge, MODS a Plain Text (one URL and tags per line). Z toho se ve dvou případech jedná o XML formáty – EndNote (XML) a MODS.

Naopak v devíti formátech je možné data exportovat. Konkrétně se jedná o formáty RIS, EndNote, BibTeX, MODS (XML), Word 2007 (Bibliography), Simple Text Citation, RSS¹²¹, RDF a Plain. Z toho se ve čtyřech případech jedná o XML formáty: MODS (XML), Word 2007 (Bibliography), RSS a RDF. Ukázky dat exportovaných v těchto XML formátech jsou k dispozici na obr. 27, 28, 29 a 30.

```
<?xml version="1.0" encoding="UTF-8"?>
<modsCollection xmlns="http://www.loc.gov/mods/v3">
<mods ID="Cassella2009">
  <titleInfo>
    <title>ECDL 2009</title>
    <subTitle>enhancing digital libraries users' experience</subTitle>
  </titleInfo>
  <name type="personal">
    <namePart type="given">Maria</namePart>
    <namePart type="family">Cassella</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">author</roleTerm>
    </role>
  </name>
  <name type="personal">
    <namePart type="given">Licia</namePart>
    <namePart type="family">Calvi</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">author</roleTerm>
    </role>
  </name>
  <originInfo>
    <issuance>monographic</issuance>
    <dateIssued>2009-11-15</dateIssued>
  </originInfo>
  <typeOfResource>text</typeOfResource>
  <genre authority="marcgt">book</genre>
</mods>
</modsCollection>
```

Obr. 27: Část záznamu exportovaného ve formátu MODS (XML)

¹²¹ RSS je zkratka výrazu *Really Simple Syndication* (opravdu jednoduchá syndikace).

```

<?xml version="1.0" encoding="UTF-8"?>
<b:Sources SelectedStyle=""
xmlns:b="http://schemas.openxmlformats.org/officeDocument/2006/bibliography"
xmlns="http://schemas.openxmlformats.org/officeDocument/2006/bibliography">
<b:Source>
  <b:Tag>Cassella2009</b:Tag>
  <b:SourceType>Book</b:SourceType>
  <b:Year>2009</b:Year>
  <b:Month>11</b:Month>
  <b:Day>15</b:Day>
  <b:Author>
    <b:Author>
      <b:NameList>
        <b:Person>
          <b>Last>Cassella</b>Last>
          <b:First>Maria</b:First>
        </b:Person>
        <b:Person>
          <b>Last>Calvi</b>Last>
          <b:First>Licia</b:First>
        </b:Person>
      </b:NameList>
    </b:Author>
  <b:Editor>
    <b:NameList> </b:NameList>
  </b:Editor>
</b:Author>
  <b>Title>ECDL 2009: enhancing digital libraries users' experience</b>Title>
</b:Source>

```

Obr. 28: Část záznamu exportovaného ve formátu Word 2007 (Bibliography)

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:prism="http://prismstandard.org/namespaces/1.2/basic/"
xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns="http://www.connotea.org/2005/01/schema#">
  <Post rdf:about="http://www.connotea.org/user/lindask/uri/60919336f04bc8d4e825d7596b6b532b">
    <title>ECDL 2009: enhancing digital libraries users' experience</title>

    <dc:subject>L. Information technology and library technology.</dc:subject>

    <dc:subject>channels.</dc:subject>

    <dc:subject>supports</dc:subject>

    <dc:subject>H. Information sources</dc:subject>

    <dc:subject>uploaded</dc:subject>

    <userBookmarkID>3546566</userBookmarkID>
    <dc:creator>lindask</dc:creator>

    <private>0</private>
  </Post>

```

Obr. 29: Část záznamu exportovaného ve formátu RDF


```

<item rdf:about="http://www.connotea.org/user/lindask/uri/60919336f04bc8d4e825d7596b6b532b">
<title>ECDL 2009: enhancing digital libraries users' experience</title>
<link>http://eprints.rclis.org/17305/</link>
<description>Posted by lindask to "L. Information technology and library technology." channels.
supports "H. Information sources" uploaded on Sat Apr 10 2010</description>
<dc:creator>lindask</dc:creator>
<dc:date>2010-04-10T18:35:24Z</dc:date>
<dc:subject>L. Information technology and library technology.</dc:subject>
<dc:subject>channels.</dc:subject>
<dc:subject>supports</dc:subject>
<dc:subject>H. Information sources</dc:subject>
<dc:subject>uploaded</dc:subject>
<slash:comments>0</slash:comments>
<content:encoded><![CDATA[<link rel="stylesheet" href="http://www.connotea.org/global.css" type="text/css" title="sty
<connotea:bookmark>http://www.connotea.org/uri/60919336f04bc8d4e825d7596b6b532b</connotea:bookmark>
<connotea:post>http://www.connotea.org/user/lindask/uri/60919336f04bc8d4e825d7596b6b532b</connotea:post>
<connotea:uri>
<dcterms:URI rdf:about="http://eprints.rclis.org/17305/">
<dc:title>ECDL 2009: enhancing digital libraries users' experience</dc:title>
<dc:creator>Maria Cassella</dc:creator>
<dc:creator>Licia Calvi</dc:creator>
<dc:date>2009-11-15</dc:date>
</dcterms:URI>
</connotea:uri>

```

Obr. 30: Část záznamu exportovaného ve formátu RSS

Interně jsou data v citačním manažeru Connotea uložena v podobě SQL databáze (konkrétně MySQL) v kompaktních tabulkách obsahujících veškeré potřebné údaje [13].

```

CREATE TABLE `citation` (
  `citation_id` int(7) unsigned NOT NULL auto_increment,
  `title` varchar(255) default NULL,
  `journal` int(7) unsigned default NULL,
  `volume` varchar(40) default NULL,
  `issue` varchar(40) default NULL,
  `start_page` varchar(40) default NULL,
  `end_page` varchar(40) default NULL,
  `pubmed` varchar(255) default NULL,
  `doi` varchar(255) default NULL,
  `asin` varchar(100) default NULL,
  `ris_type` varchar(4) default NULL,
  `raw_date` varchar(40) default NULL,
  `date` date default NULL,
  `last_modified_date` date default NULL,
  `user_supplied` int(1) unsigned NOT NULL default '0',
  `cs_module` varchar(40) default NULL,
  `cs_type` varchar(255) default NULL,
  `cs_source` varchar(255) default NULL,
  `created` datetime NOT NULL default '0000-00-00 00:00:00',
  PRIMARY KEY (`citation_id`),

```

Obr. 31: Část dotazu v jazyce SQL pro vytvoření tabulky „citation“

Pokud se týká automatického přebírání dat do manažeru Connotea, existují dva způsoby – prostřednictvím tlačítka či záložky (*bookmarklet*) „Add to Connotea“

v internetovém prohlížeči nebo zadáním identifikátoru DOI do pole pro URL ve webovém formuláři. Pokud uživatel zvýrazní na webové stránce DOI a použije tlačítko v prohlížeči, Connotea rovněž získá příslušná data s využitím DOI. Do tohoto citačního manažeru mohou být ukládány jakákoliv URL, spolehlivé automatické přebírání dat je umožněno pouze ze zdrojů, s nimiž Connotea umí přímo spolupracovat (např. Nature.com, PubMed, Science či PloS). Je také možné získávat data automaticky prostřednictvím kanálů RSS. Vždy je však potřeba, aby daný dokument měl URL (údaje o dokumentech bez URL nelze do Connotey ukládat) [215].

5.4 Citace 2.0

Citační manažer Citace 2.0 vznikl (pod názvem Citace.com) v roce 2004, verze 2.0 pak byla spuštěna v roce 2009. Na jeho vývoji se podílí zejména Martin Krčál z Ústřední knihovny Fakulty sociálních studií Masarykovy univerzity. Citační manažer získal již dvakrát Cenu Infora – v letech 2005 a 2010 [216].

Na rozdíl od ostatních popisovaných citačních manažerů je tento nástroj zaměřen na generování záznamů v souladu s normami ČSN ISO 690 a 690-2 (možnost vytvořit záznamy v souladu s těmito normami není u jiných citačních manažerů často nabízena).

Pokud se týká importu dat, umožňuje nástroj Citace 2.0 v současné době přebírání záznamů od jiných svých uživatelů nebo z externích zdrojů (vybraných knihovních katalogů a databáze knih internetového knihkupectví Amazon). Z knihovních katalogů jsou data přebírána prostřednictvím protokolu SRU¹²², z databáze knihkupectví Amazon prostřednictvím API¹²³ v XML formátu [217, s. 99-100]. Import z těchto externích zdrojů se tedy týká pouze knih s přiděleným identifikátorem ISBN.¹²⁴ Tuto skutečnost názorně dokládá obr. 32.

¹²² SRU je zkratka výrazu *Search/Retrieval via URL* (vyhledávání pomocí URL).

¹²³ API je zkratka výrazu *Application Programming Interface* (aplikační programovací rozhraní).

¹²⁴ Do budoucna je však počítáno s rozšířením i na záznamy jiných typů dokumentů, zejména časopiseckých článků [217, s. 101].

Import citace

Toto [místo](#) slouží pro import již vytvořených citací, případně umožňuje vytvořit citaci z údajů v externích [katalozích](#) (v současnosti katalog Library of Congress, The Britttish Library, Masarykovy univerzity a z [databáze](#) knih Amazonu).

Hledat na Citace.com

Název

Příjmení

ISBN

Hledat v jiných zdrojích

ISBN

Záznamy nalezené na Citace.com

CEJPEK, Jiří. *Informace, komunikace a myšlení : úvod do informační vědy*. 2. Praha : Karolinum, 2005. 233 s. ISBN 80-246-1037-X. (**kniha**, použito 1 x) [[převzít](#)]

CEJPEK, Jiří. *Dějiny knihoven a knihovnictví*. 2., dopl. vyd. Praha : Karolinum, 2002. 247 s. ISBN 80-246-0323-3. (**kniha**, použito 1 x) [[převzít](#)]

Obr. 32: Import záznamů do citačního manažeru Citace.com

Jiný způsob importu dat (např. z lokálně uložených souborů ve formátu RIS, BibTeX apod.) zatím nabízen není.

Data je možné exportovat ve formátech HTML, RTF, BibTeX a TXT. Ve formátech HTML, RTF a TXT jsou data exportována v podobě záznamu v souladu s výše uvedenými normami (záznam TXT však pochopitelně neobsahuje formátování), pouze ve formátu BibTeX jsou data k dispozici ve strukturované podobě.

Lze tedy konstatovat, že pro import ani export dat nejsou využívány XML formáty.

Interní formát uložení dat není zveřejněn (nejedná se o software s otevřeným zdrojovým kódem).

5.5 CiteULike

Citační manažer CiteULike¹²⁵ [14], vyvinul Richard Cameron v roce 2004. Následně byla vytvořena společnost Oversity Ltd., která službu spravuje a rozvíjí. Službu CiteULike podporuje i vydavatelství Springer.

Citační manažer CiteULike umožňuje importovat data ve formátech RIS a BibTeX. Export je možné realizovat v těchto formátech (záznamy je také možné exportovat do souborů ve formátech PDF, RTF či ve formě čistého textu). Ani v jednom případě se tedy nejedná o XML formát.

CiteULike není software s otevřeným zdrojovým kódem, k uložení dat je využívána databáze PostgreSQL [218].

Pokud se týká automatického přebírání dat z jiných zdrojů, pro uživatele nejjednodušší si do internetového prohlížeče nainstalovat tlačítko či záložku (*bookmarklet*) a jeho prostřednictvím pak přímo importovat data; alternativou může být využití upravených bran protokolu (*custom protocol handlers*) [219]. Příklad záznamu automaticky importovaného z webového sídla časopisu Nature do citačního manažeru CiteULike prostřednictvím bookmarkletu je k dispozici na obr. 33.

¹²⁵ Označení verzí není veřejně používáno, lze pouze předpokládat interní označování verzí vývojáři.

CiteULike MyCiteULike Search Logged in as lindask

New article: where would you like to file it?

Let us know where, and how, you want this filed.

Title: Ecological interactions are evolutionarily conserved across the entire tree of life

Published: 2010-06-02

Abstract: Ecological interactions are crucial to understanding both the ecology and the evolution of organisms^{1, 2}. Because the phenotypic traits regulating species interactions are largely a...

Authors: Gomez JM, Verdu M, Perfectti F

Tags:

Note: This is just a list of keywords which you'd like to associate with the article. Prefix private tags with "*". You can also mark articles for the attention of one of your [connections](#) as "for:username", but only in personal libraries, not groups.

Post to: Your library

Priority:

Privacy: Don't let other people see that I've posted this article

Notes:

Attachment:

Keep original filename (If unchecked, filename will be based on author-year-title, but you can change this later.)

Obr. 33: Import záznamů do citačního manažeru CiteULike

V této souvislosti je potřeba poznamenat, že CiteULike rozlišuje mezi tzv. důvěryhodnými daty (*trusted data*) a daty nedůvěryhodnými (*untrusted data*). Za důvěryhodná data jsou pokládána ta, která jsou z tzv. schválených webových sídel (*approved sites*), jejichž prostřednictvím jsou publikovány články v recenzovaných časopisech [220]. Dále je možné záznamy vyhledávat v databázi CiteULike, prostřednictvím identifikátoru DOI či PMID.

Lze přebírat i data přímo ze služby CiteULike, a to do softwarů BibDesk, Mendeley, popř. jakýchkoliv HTTP¹²⁶ klientů. Je také možné pracovat se záznamy ve formátu COinS, získávat je prostřednictvím kanálů RSS či formátu JSON¹²⁷. Užitečná je také nabídka exportu dat do systému Delicious [220].

¹²⁶ HTTP je zkratka výrazu *HyperText Transfer Protocol* (přenosový protokol hypertextu).

¹²⁷ JSON je zkratka výrazu *JavaScript Object Notation* (JavaScriptový objektový zápis).

5.6 Dílčí závěry

Citační manažery začínají hrát velmi podstatnou úlohu při přípravě vědeckých a jiných odborných textů. V posledních letech se stále více prosazují citační manažery dostupné formou online služby. Často se zároveň jedná o nástroje s otevřeným zdrojovým kódem.

Mezi tyto nové typy citačních manažerů se řadí například nástroje Zotero, Connotea, Citace 2.0 a CiteULike. S výjimkou softwaru Citace 2.0 nabízejí vždy alespoň dva formáty pro import dat (s nejširší nabídkou formátů se setkáváme u citačního manažeru Connotea). Ze všech čtyř vybraných nástrojů potom mohou být data exportována v minimálně dvou formátech (nejvyšší počet exportních formátů má Connotea).

Tři XML formáty pro import a export dat jsou k dispozici u softwaru Zotero, v případě Connotey jsou dostupné čtyři XML formáty pro export dat. Nástroje Citace 2.0 a CiteULike naopak XML formáty nepodporují.

Pokud se týká interního uložení dat, jsou ve třech případech používány SQL databáze, způsob uložení dat ve čtvrtém zkoumaném citačním manažeru Citace 2.0 není zveřejněn.

Všechny čtyři nástroje potom umožňují automatické přebírání citačních dat z externích zdrojů, například pomocí tlačítek v prohlížeči nebo zadání URL, DOI či ISBN.

Automatizace práce s citačními daty s využitím webových nástrojů patří mezi dynamicky se rozvíjející oblasti. To potvrzuje i skutečnost, že přes nedávný vznik všech čtyř nástrojů (v rozmezí let 2004-2006) získaly tyto nástroje již značnou oblibu a široký okruh uživatelů.

6. Typologie citovaných dokumentů

6.1 Úvod

Tato kapitola se nejprve zabývá samotným termínem „dokument“, resp. pojmem, který tento termín označuje, a snaží se postihnout významné změny, k nimž v tomto směru došlo s rozvojem elektronického prostředí. Dále prezentuje přehled vybraných zdrojů věnovaných typologiím dokumentů. Jádrem kapitoly tvoří části věnované typologiím dokumentů, s nimiž pracují citační formáty, styly a ontologie, jimiž jsme se zabývali v předchozích částech práce. Doplnkově jsou zahrnuty vybrané typologie dokumentů, které jsou používány v oblasti knihovnictví. Nejedná se tedy o taxativní výčet typologií, ale o charakteristiku jejich určitého vzorku.¹²⁸ Přednost byla dána typologiím užívaným v mezinárodním měřítku,¹²⁹ proto jsou také názvy jednotlivých typů dokumentů uváděny v anglickém jazyce.

U konkrétních typologií byly hledány zejména odpovědi na následující otázky:

- Jak probíhá aktualizace typologie? Je typologie postupně rozšiřována?
- Je přehled všech typů dokumentů k dispozici na jednom místě nebo jsou potřebné údaje rozptýleny na více místech?
- Lze typologii pokládat za jednoznačnou? Jsou dostupné podrobnější vysvětlivky týkající se očekávaného užití typů dokumentů v praxi?

V případě citačních formátů RIS, ISI, RefWorks a BibTeX je navíc zahrnuta ukázka možného mapování položek z typologií obsažených v těchto formátech (vždy ve vztahu k předchozím zmiňovaným formátům).

6.2 Termín „dokument“

Zabýváme-li se typologií dokumentů, je na místě stručně vyložit obsah a rozsah pojmu označovaného v české odborné terminologii z oblasti informační vědy jako dokument.

¹²⁸ Z těchto důvodů není zahrnuta například typologie v metadatovém schématu Dublinské jádro či typologie MIME, o nichž podrobněji pojednává E. Bratková [222, s. 17-24], nebo typologie, kterou lze extrahovat z normy (ČSN) ISO 5127 [223].

¹²⁹ Není například zahrnuta typologie formálních deskriptorů, které tvoří součást věcných autorit spravovaných Národní knihovnou ČR [221].

V České terminologické databázi knihovnictví a informační vědy se setkáme s výkladem, že dokumentem je míněn „[i]nformační pramen tvořený nosičem informací a množinou informací na něm fixovaných a sloužící k přenosu dat v čase a prostoru“ [41, heslo dokument]. M. Nováková [224, s. 32] uvádí definici, která říká, že „[d]okument je prostředek sociální komunikace, tvořený množinou materiálně fixovaných a uspořádaných dat nebo informací určených pro jejich přenos v čase a prostoru“. Podle této autorky jsou dokumenty podřazeny hmotným informačním pramenům a ty informačním pramenům obecně [224, s. 32].

S ohledem na vznik sítově dostupných dokumentů se však již před vznikem této databáze začaly i v českých odborných časopisech objevovat články upozorňující na skutečnost, že termín dokument již v tomto novém kontextu nelze pokládat za adekvátní [225], [226], [227, s. 56]. Zaměříme-li se na elektronické prostředí, je možné uvažovat o užití termínu (resp. terminologického sousloví) nedokumentový informační pramen¹³⁰ (nebo zdroj) [225]. Pojmeme-li naopak otázku šířeji, tj. neomezíme-li se na elektronické prostředí, může se jako vhodný jevit obecnější termín informační objekt [226].

Navíc je s rozvojem elektronického publikování nutné diskutovat o aktuálních hranicích dokumentu či informačního objektu (viz např. [228] či [229]) – např. kam až sahá elektronická monografie obsahující odkazy na další uzly sítě? Jsou její součástí odkazované dokumenty, které jsou uloženy na stejném serveru, nebo i odkazy směřující na dokumenty nacházející se na jiném serveru (fyzickém místě)? Z hlediska uživatele sice není lokalizace konkrétních částí dokumentu příliš podstatná (důležité je získání hledané informace, nikoliv pouze informace o existenci dokumentu, v němž by mohla být obsažena odpověď na jeho otázku), z hlediska zpracovatelů informací (informačních pracovníků, knihovníků) však vystupuje do popředí spor obsah (*content*) versus nosič (*carrier*).

V této práci je zjednodušeně (s využitím výše zmíněného výkladu z České terminologické databáze knihovnictví a informační vědy) za dokument pokládán informační pramen, který existuje v hmotné nebo nehmotné podobě, obsahuje množinu informací a slouží k přenosu těchto informací v čase a prostoru. Ještě

¹³⁰ Výraz pramen je v knihovnicko-informační praxi užíván v širším, resp. jiném smyslu než např. v oblastí historie.

obecněji lze jako dokument označit grafický (nejenom textový) záznam potenciálních informací¹³¹. Jako ekvivalent termínu „dokument“ je v této práci užíván výraz „zdroj“.

V návrhu experimentálního XML formátu pro citační data je také – přes výše uvedené skutečnosti, které problematizují vhodnost jeho užití – tento termín (resp. termín „document“) upřednostněn. Důvod je pragmatický – jedná se totiž o stručné a tradičně užívané označení tohoto pojmu, byť již dnes nevystihuje všechny jeho rysy.

Podrobné pojednání o terminologických otázkách nepatří mezi cíle předkládané disertační práce, proto je tato problematika uvedena pouze stručně.

6.3 Přehled vybraných zdrojů věnovaných typologiím dokumentů

V této podkapitole jsou stručně přiblíženy vybrané zdroje týkající se typologie dokumentů, zejména pak díla informačních vědců a odborníků z příbuzných oblastí. Ve vybraných případech zachycena kritéria členění dokumentů u daných autorů. Bez znalosti těchto kritérií totiž může nastat situace, kdy typologie bude ve výsledku vypadat podobně či stejně jako jiná typologie vytvořená s uplatněním jiných kritérií.¹³²

Úvodem je potřeba předeslat, že se zpracováním různých typologií dokumentů nebo souvisejících dílčích otázek se setkáme u řady významných autorů. Např. P. Otlet [231] ve svém díle *Traité de documentation* uvažuje dokonce nejenom o textových dokumentech, ale také o řadě jiných typů dokumentů (například přírodních objektech či modelech) [231, zejm. s. 217], [232, s. 355], [233].

S typologií dokumentů se setkáme rovněž v *Návrhu na triedenie knižničnej vedy* J. Millse, který je dostupný ve slovenském překladu (spolu s textem B. C. Vickeryho *Fazetová klasifikácia*) [234]. Jedná se o jeden z výstupů práce Výzkumné skupiny pro třídění (*Classification Research Group*, CRG [235]). Je v něm mj.

¹³¹ Výjimečně v praxi dochází k určitému sporu, kdy je termín užíván i v souvislosti s údaji týkajícími se nezaznamenaného poznání či nezaznamenaných informací. Existují však i časopisy, které v pokynech pro autory píšící články explicitně zakazují citování osobních sdělení či obecně nepublikovaných informací.

¹³² Podobně J. Searle naznačuje, že věci se mohou dít na základě různých pravidel. Uvádí příklad svých dětí, které se učily parkovat a řídily se přitom určitým pravidlem. On sám se tímto pravidlem však neřídil, ale z hlediska vnějšího pozorovatele byl v obou případech způsob parkování totožný [230, s. 50-51].

zdůrazněno, že členění bibliografické, resp. literární formy se může velmi lišit podle oboru [234, s. 29]. V konkrétním případě knihovní vědy jsou mezi všeobecnými pomocnými znaky formy prostřednictvím kategorií (faset) uvedena tato hlediska: „forma podání“, „fyzická forma tiskoviny“, „způsob psaní“, „typ symbolů“, „základní materiál“, „smyslový kanál“, „podle původu, původ díla“, „podle činitelů týkajících se vydání“ [234, s. 73-75].

Typologií dokumentů, resp. informačních zdrojů a prostředků, a vztahy mezi dokumenty vyjádřenými formou grafů se zabývá i J. Freytag. Podle umístění dokumentů v grafech rozlišuje jednotlivé typy dokumentů primárních (nejsou koncovými uzly grafu) a sekundárních (naopak jsou koncovými uzly grafu) [236].

G. Rózsa nastiňuje možná kritéria členění dokumentů z oblasti sociálních věd, konkrétně v ekonomii. Např. u periodik pracuje se čtyřmi kritérii – rozlišuje typy periodik podle oblasti, periodicity, vydavatele a typu [237].

B. C. Vickery a A. Vickeryová se ve své práci „Information Science in Theory and Practice“ věnují mj. komunikačním médiím a jejich třídění, resp. typologii. Za podnětný lze v tomto případě pokládat původně Duncanův přehled médií uspořádaných podle rostoucích jednotkových nákladů, obtížnosti pořízení a potenciální velikosti publika (přehled začíná rukopisnými poznámkami přednášejícího nebo účastníka a končí televizními přenosy). Autoři zmiňují také rozdíly mezi komunikačními médii ze šesti různých pohledů, které identifikoval Schramm (podle smyslů, jichž se to týká, podle možnosti zpětné vazby, podle míry možnosti řízení ze strany příjemce, podle typu zakódování zprávy, podle multiplikativní síly a podle síly uchování zprávy) [238, s. 49-50].

D. Kaye se na typologii dokumentů dívá z hlediska formátu, statusu a umístění [239].

Podnětným způsobem k tématu typologie dokumentů (konkrétně dokumentů webových) přistupuje M. A. Rosso, který se ve svém výzkumu zabýval identifikací typů webových dokumentů (označovaných ovšem jako žánry) uživateli [240].

Ve slovenském prostředí se na problematiku typologie dokumentů z pohledu informační vědy¹³³ dlouhodobě zaměřovala M. Nováková [241], [242], [224].

¹³³ Dříve byl častěji používán termín „informatika“.

V českém prostředí se těmito otázkami zabývaly mj. H. Vodičková [243], H. Vítková [244] a E. Bratková [225]. Stručnou charakteristiku kritérií, podle nichž lze členit dokumenty, podává i M. Baslová-Randová ve své diplomové práci [245, zejm. s. 17-21].

S názvem typologie dokumentů, resp. kritérií členění dokumentů, se setkáme i v České terminologické databázi knihovnictví a informační vědy [41, heslo dokument]. V rámci hesla dokument je uvedeno, že „[d]okumenty se dělí podle řady kritérií, např. podle způsobu záznamu dat (písemné, obrazové, zvukové, audiovizuální, strojem čitelné – elektronické či digitální), podle odvozenosti obsahu (primární, sekundární a terciární), podle kontinuity (periodické a neperiodické), podle stupně zveřejnění (zveřejněné, nezveřejněné, interní).“

Typologii dokumentů je navíc věnována pozornost v řadě studijních textů. Např. J. Skolek stručně charakterizuje typy dokumentů s ohledem na typ informací, který obsahují (od knih jako klasického prototypu dokumentu až např. po separáty a preprinty či zvukové, obrazové a audiovizuální dokumenty); doplňuje však také, že dokumenty je možné členit i podle jiných kritérií (jako příklad je uveden primární dokument, sekundární dokument a odvozený neboli sumarizační dokument) [246, s. 18-21].

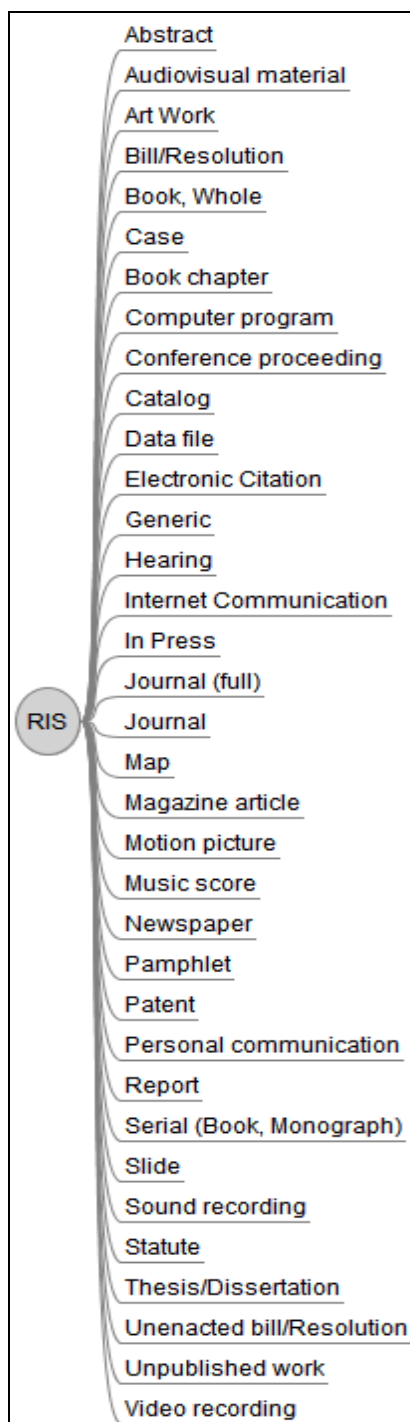
B. Sedláčková se zaměřuje na textové odborné dokumenty a uvádí členění, s nímž se především v knihovnicko-informační praxi pracuje jako s celky [248, s. 24-53]¹³⁴. Pracuje se čtyřmi třídami (knižní publikace, seriálové publikace, speciální publikace a nepublikované dokumenty) a v jejich rámci s typy (konkrétně v případě knižních publikací se jedná o monografie, sborníky, učebnice, příručky, encyklopedie a slovníky) [248, s. 24]. Zmiňuje rovněž další třídy dokumentů vymezené podle vyjadřovacích prostředků, a to třídy dokumentů obrazových, zvukových a audiovizuálních [248, s. 51], v samostatné kapitole navíc pojednává o informačních technologiích v dokumentové komunikaci [248, s. 54-61].

J. Šilhánek preferuje kategorizaci dokumentů (resp. informačních zdrojů) podle jejich roli v procesu vzniku a komunikace odborných informací. Informační zdroje člení na primární, sekundární (systematické) a terciární [249, s. 19-29].

¹³⁴ Navazuje tím na svou předchozí práci [247].

6.4 Typologie dokumentů v citačních formátech

6.4.1 RIS



Obr. 34: Typologie dokumentů ve formátu RIS

Formát RIS nabízí celkem 35 různých typů dokumentů (viz obr. 34).

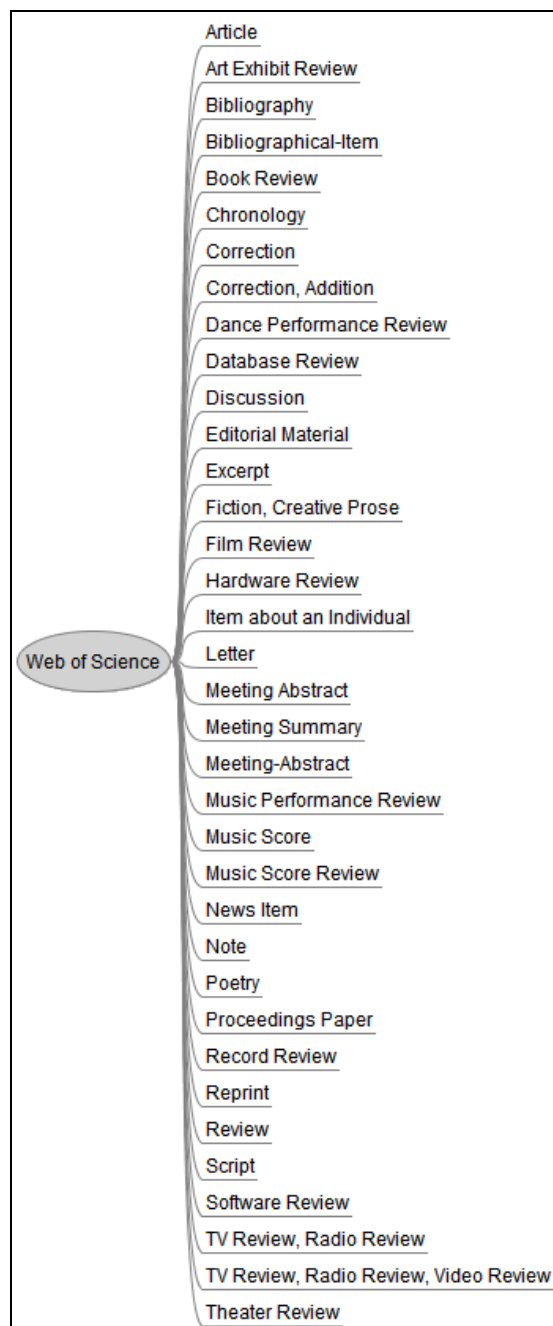
Jednotlivé položky nejsou hierarchicky uspořádány – všechny typy dokumentů se tak nacházejí na stejné úrovni, např. typ „Book, Whole“ (celá kniha) a typ „Book chapter“ (kapitola knihy), mezi nimiž jinak existuje vztah celek-část. Značnou nevýhodou typologie užití ve formátu RIS je skutečnost, že k jednotlivým položkám nejsou k dispozici vysvětlivky (svým způsobem jsou však vysvětlivkami již výše uvedená označení typů, neboť z hlediska formátu je podstatné zkrácené vyjádření typu dokumentu formou názvu pole, tj. např. „ABST“, „ADVS“ či „ART“ v případě prvních třech položek v pořadí). Navíc není zcela zřejmé, jakým způsobem je vhodné zařazovat dokumenty odpovídající více charakteristikám, např. zda by článek v online časopise měl mít přiřazen typ „Electronic Citation“ (elektronická citace), „Journal“ (časopis) či „Journal (full)“ (úplný název časopisu). To může vyústit v preferenci typu „Generic“ (obecný). Jako problematický typ dokumentu je možné též vnímat typ „In Press“ (v tisku) – jedná se totiž o status dokumentu, vlastním dokumentem pak může být např. časopisecký článek či příspěvek ve sborníku.

Typologie je dostupná na jedné webové stránce. Aktualizace je pravděpodobně prováděna spíše sporadicky, zdrojová webová stránka byla naposledy aktualizována v roce 2001 [15].

6.4.2 ISI

Vzhledem k nedostupnosti specifikace formátu ISI (viz výše) byl v tomto případě proveden průzkum typologie dokumentů, s nímž pracují databáze dostupné prostřednictvím v rámci systému Web of Science. Průzkum vycházel z předpokladu, že může existovat vztah mezi typologií užívanou v tomto systému a typologií dokumentů ve formátu ISI.

Z nápovědy k systému Web of Science [16] bylo zjištěno, že je používána typologie dokumentů, již tvoří celkem 36 položek (viz obr. 35).



Obr. 35: Typologie dokumentů v systému Web of Science

V systému Web of Science byly následně provedeny rešerše s cílem získání záznamů jednotlivých typů dokumentů ve formátu ISI. Příklad znění rešeršního dotazu zadaného v rozhraní pro pokročilé vyhledávání je následující: TS=(digital libraries), resp. TS=(digital libraries) AND Document Type=(Article). Typy dokumentů byly vybírány z menu (tj. jednalo se o filtr). Znění uživatelem zadaného

dotazu bylo rešeršním systémem následně automaticky doplněno o typ dokumentu. S ohledem na cíl průzkumu byl v případě rešerší bez výsledků upraven dotaz (v případě potřeby opakovaně) tak, aby vždy byla získána sada s alespoň jedním výsledkem.

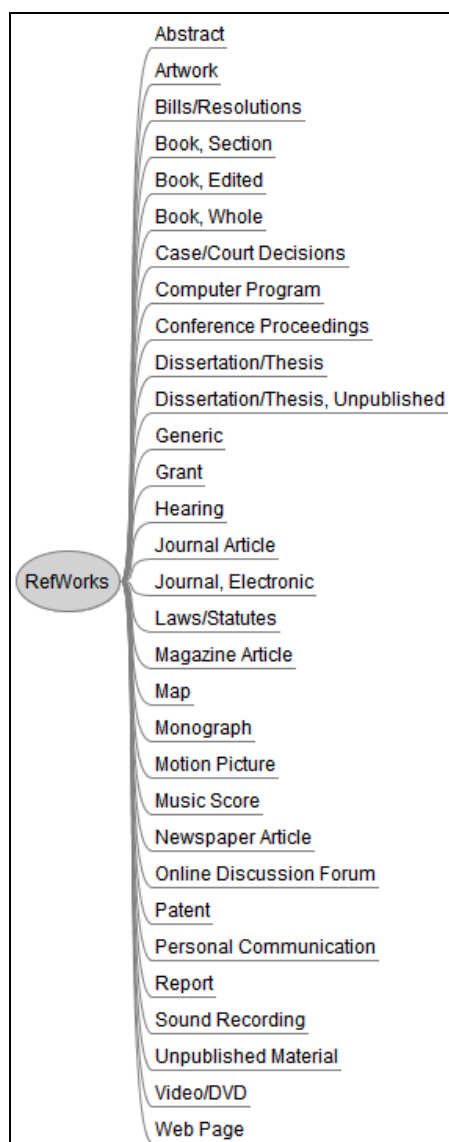
Při průzkumu bylo získáno celkem 33 souborů záznamů ve formátu ISI (v každém souboru byl obsažen právě jeden záznam). Ve třech případech bylo zjištěno, že ačkoliv jsou v nápovědě systému Web of Science i v roletkovém menu rešeršního rozhraní některé položky uvedeny jako samostatné, systém je pokládá za synonyma, resp. o druhou obdobnou položku automaticky rozšiřuje dotaz. Jednalo se o tři dvojice položek: 1) „Bibliographical-Item“ a „Item About An Individual“, 2) „Correction“ a „Correction, Addition“ a 3) „TV Review, Radio Review“ a „TV Review, Radio Review, Video Review“. Tyto dvojice položek byly v exportovaných souborech zastoupeny vždy pouze jednou.

Při studiu 33 souborů bylo zjištěno, že v poli s názvem „PT“, do něhož je ve formátu ISI zapisován typ dokumentu, byla ve všech 33 případech uvedena hodnota „J“, kterou lze považovat za zkratku výrazu „Journal“ (časopis). Bylo tomu tak i v případě, že podle typologie systému Web of Science se jednalo o typ dokumentu „Proceedings Paper“ (příspěvek v konferenčním sborníku).

Úvodní předpoklad existence vztahu mezi typologií užívanou v systému Web of Science a typologií dokumentů ve formátu ISI se tedy nepotvrdil. Lze konstatovat, že pro náš účel není typologie dokumentů ve formátu ISI dále použitelná, resp. inspirativní, neboť se nepodařilo zjistit existenci jiného typu dokumentu než časopisu, resp. časopiseckého článku.

6.4.3 RefWorks

Typologie dokumentů použitá ve formátu RefWorks [17] je složena z celkem 31 položky (viz obr. 36).



Obr. 36: Typologie dokumentů ve formátu RefWorks

Typologie je velmi podobná typologii užívané ve formátu RIS – v deseti případech jsou v obou typologiích použity zcela shodné názvy prvků („Abstract“, „Book, Whole“, „Generic“, „Hearing“, „Magazine Article“, „Motion Picture“, „Music Score“, „Patent“, „Report“ a „Sound Recording“). Ve dvanácti případech jsou pak názvy položek velmi podobné, resp. liší se v drobnostech, viz tab. 40.

Tab. 40: Typy dokumentů ve formátech RIS a RefWorks

Typ dokumentu ve formátu RIS	Odpovídající typ dokumentu ve formátu RefWorks
Art Work	Artwork
Bill/Resolution	Bills/Resolutions
Book chapter	Book, Section
Case	Case/Court Decisions
Computer program	Computer Program
Conference proceeding	Conference Proceedings
Thesis/Dissertation	Dissertation/Thesis
Journal (nebo Journal (full))	Journal Article
Newspaper	Newspaper Article
Personal communication	Personal Communication
Unpublished work	Unpublished Material
Video recording	Video/DVD

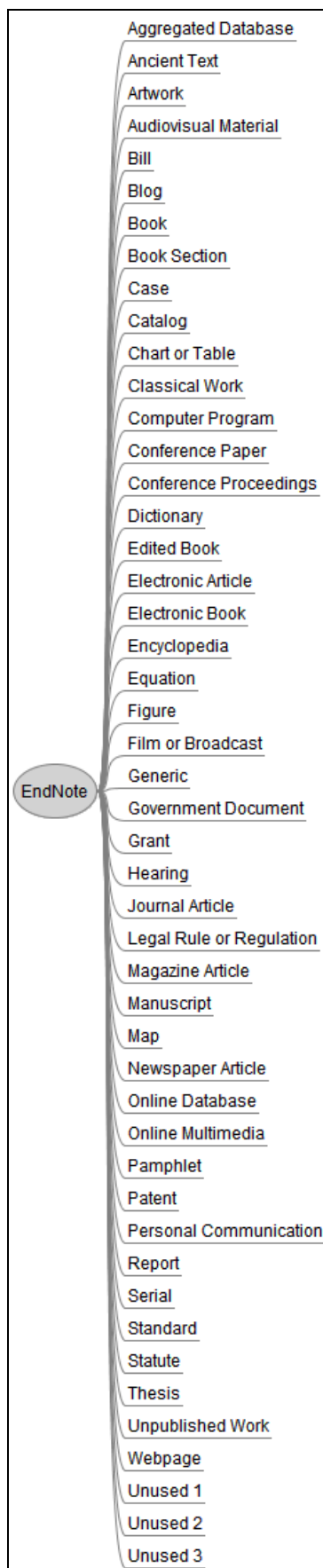
I v případě typů dokumentů ve formátu RefWorks lze konstatovat, že při její tvorbě byly uplatněny různé pohledy. Ve specifikaci formátu není uvedeno, jakým způsobem vybrat typ dokumentu v případě, že pro daný dokument by se jako vhodný typ jevílo více položek ze seznamu (vysvětlivky týkající se způsobu užití položek nejsou ve specifikaci k dispozici). Z příkladů typů „Dissertation/Thesis, Unpublished“ a „Unpublished Material“, resp. „Journal, Electronic“ a „Web Page“ lze pouze nepřímou dovést, že přednost má specifitější typ, tj. typ, který daný dokument charakterizuje přesněji.

Lze také konstatovat, že typologie nevyužívá hierarchii, proto se na stejné úrovni objevují položky, které mezi sebou reálně mají vztah celek-část (např. „Book, Section“ je částí „Book, Whole“).

Typologie je dostupná na jedné webové stránce, na níž je umístěna celá specifikace formátu RefWorks [17]. Na ní není uvedeno datum poslední aktualizace, lze pouze předpokládat, že aktualizace je prováděna průběžně.

6.4.4 EndNote

Typologie dokumentů, s níž pracuje formát EndNote, je složena z celkem 48 položek (viz obr. 37), z nichž poslední tři jsou označeny jako nepoužívané („Unused



Obr. 37: Typologie dokumentů ve formátu EndNote

1“, „Unused 2“ a „Unused 3“). V manuálu k softwaru Endnote [18, s. 711-739 (Appendix C Reference Types and Fields)] je navíc u každého typu dokumentu k dispozici přehled polí, které jsou u daného typu dostupné v softwaru EndNote.

Porovnáme-li typologii dokumentů ve formátu EndNote s typologií dokumentů ve formátu RIS, zjistíme, že v jedenácti případech se setkáme se zcela shodnými názvy položek („Audiovisual Material“, „Case“, „Catalog“, „Generic“, „Hearing“, „Magazine Article“, „Map“, „Pamphlet“, „Patent“, „Report“ a „Statute“). Ve dvanácti případech jsou pak názvy položek obdobné (viz tab. 41).

Tab. 41: Typy dokumentů ve formátech RIS a EndNote

Typ dokumentu ve formátu RIS	Odpovídající typ dokumentu ve formátu EndNote
Art Work	Artwork
Bill/Resolution	Bill
Book, Whole	Book
Book chapter	Book Section
Computer program	Computer Program
Conference proceeding	Conference Proceedings
Journal (nebo Journal (full))	Journal Article
Newspaper	Newspaper Article
Personal communication	Personal Communication
Serial (Book, Monograph)	Serial
Thesis/Dissertation	Thesis
Unpublished work	Unpublished Work

Porovnáme-li typologii dokumentů ve formátu EndNote s typologií dokumentu ve formátu RefWorks, dospějeme k závěru, že ve třinácti případech jsou v obou typologiích použity stejné názvy prvků („Artwork“, „Computer Program“, „Conference Proceedings“, „Generic“, „Grant“, „Hearing“, „Journal Article“, „Magazine Article“, „Map“, „Newspaper Article“, „Patent“, „Personal Communication“ a „Report“). V osmi případech jsou potom názvy prvků obdobné (viz tab. 42).

Tab. 42: Typy dokumentů ve formátech RefWorks a EndNote

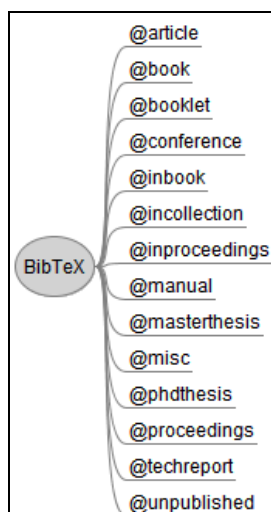
Typ dokumentu ve formátu RefWorks	Odpovídající typ dokumentu ve formátu EndNote
Bills/Resolutions	Bill
Book, Whole	Book
Case/Court Decisions	Case
Journal, Electronic	Electronic Article
Book, Edited	Edited Book
Motion Picture	Film or Broadcast
Dissertation/Thesis	Thesis
Unpublished Material	Unpublished Work

Z typologie dokumentů ve formátu EndNote je – stejně jako u předchozích typologií – zřejmé mísení pohledů a nevyužití hierarchických vazeb. Na druhou stranu je díky výčtu polí, která lze u daného typu dokumentu v softwaru EndNote použít, možné v konkrétních příkladech přesněji odhadnout vhodnost typu dokumentu. Tato skutečnost však k jednoznačnosti typologie může pouze přispívat.

Na závěr poznamenejme, že aktualizace typologie může být prováděna v souvislosti s aktualizací softwaru EndNote a souvisejících manuálů.

6.4.5 BibTeX

Formát BibTeX používá ve srovnání s výše rozebíranými formáty stručnou typologii dokumentů – je složena z pouhých čtrnácti položek (viz obr. 38), z nichž lze ovšem dvě („@inproceedings“ a „@conference“) v podstatě považovat za shodné [19, s. 14-15]. V dokumentaci formátu BibTeX a v dalších dokumentech pojednávajících o tomto formátu (viz např. i [19]) jsou často k dispozici vysvětlivky



Obr. 38: Typologie dokumentů ve formátu BibTeX

týkající se způsobu užití jednotlivých typů dokumentů v praxi. Typologie jako celek je pak typicky dostupná na jednom místě v rámci daného dokumentu o formátu.

Vzhledem k nízkému počtu prvků je porovnání typologie dokumentů ve formátu BibTeX realizováno formou mapovací tabulky (viz tab. 43). U některých položek je mapování velmi přibližné, popř. je naznačeno více možných řešení (např. typ dokumentu „@manual“). V některých případech odpovídá různým typům dokumentu ve formátu BibTeX jeden typ dokumentu ostatních formátů (např. typy dokumentu „@inproceedings“ a „@conference“).

Tab. 43: Typy dokumentů ve formátech BibTeX, RIS, RefWorks a EndNote

Typ dokumentu ve formátu BibTeX	Odpovídající typ ve formátu RIS	Odpovídající typ ve formátu RefWorks	Odpovídající typ ve formátu EndNote
@article	Journal (full) nebo Journal	Journal Article	Journal Article
@book	Book, Whole	Book, Whole	Book
@booklet	Pamphlet	Book, Whole nebo Generic	Pamphlet
@conference	Conference proceeding	Conference Proceedings	Conference Proceedings
@inbook	Book chapter	Book, Section	Book Section
@incollection	Book chapter	Book, Edited	Edited Book
@inproceedings	Conference proceeding	Conference Proceedings	Conference Paper nebo Conference Proceedings
@manual	Book, Whole, Pamphlet nebo Generic	Book, Whole nebo Generic	Book, Pamphlet nebo Generic
@masterthesis	Thesis/Dissertation	Dissertation/Thesis	Thesis
@mics	Generic	Generic	Generic
@phdthesis	Thesis/Dissertation	Dissertation/Thesis	Thesis
@proceedings	Conference proceeding	Conference Proceedings	Conference Proceedings
@techreport	Report	Report	Report
@unpublished	Unpublished work	Unpublished Material	Unpublished Work

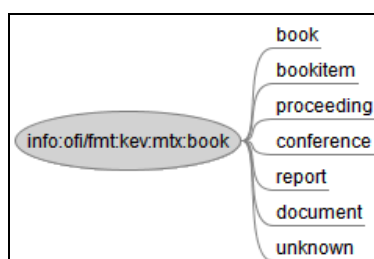
Formát BibTeX v typologii dokumentů rovněž nevyužívá hierarchii; z hlediska konstrukce typologie (resp. výsledku tohoto úsilí) však lze konstatovat relativně vysokou míru konzistence přístupu. K tomu jistě přispívají i nízký počet prvků v typologii, a dále též skutečnost, že oficiální typologie pochází již z 80. let 20.

století, kdy ještě řada typů online dokumentů nepatřila mezi běžně se vyskytující typy dokumentů. To je však zároveň možné považovat za negativní rys, neboť nové typy dokumentů (zejména online dokumenty) lze používat pouze neoficiálně.

6.4.6 OpenURL COinS

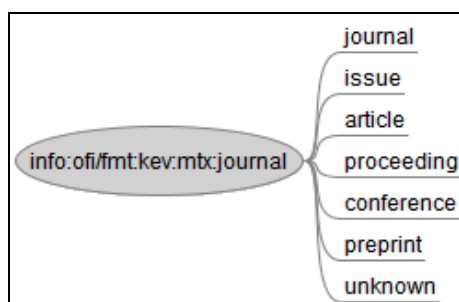
V případě formátu OpenURL COinS záleží typologie dokumentů na použitém metadatovém formátu.

Jako příklad můžeme (v návaznosti na předchozí pasáž věnovanou tomuto formátu) uvést metadatový formát `info:ofi/fmt:kev:mtx:book` [20] určený pro údaje o knihách. V tomto schématu je k dispozici prvek „genre“, který může nabývat celkem sedmi hodnot (viz obr. 39).



Obr. 39: Typologie knih a jejich částí ve formátu `info:ofi/fmt:kev:mtx:book`

Formát `info:ofi/fmt:kev:mtx:journal` [250] určený pro časopisecké a příbuzné zdroje rovněž pracuje se sedmi hodnotami typů v prvku „genre“ (viz obr. 40).



Obr. 40: Typologie časopiseckých a příbuzných zdrojů ve formátu `info:ofi/fmt:kev:mtx:journal`

V obou případech je u každé položky doplněna vysvětlivka upřesňující očekávaný způsob použití položky, tj. způsob užití typologií by měl být jednoznačný.

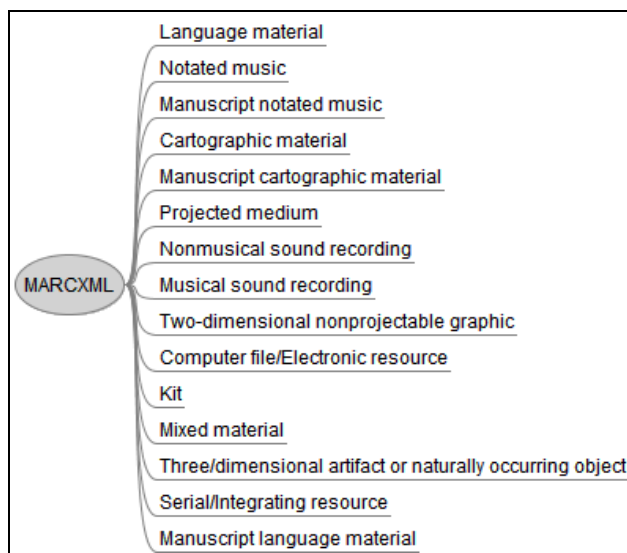
Obě typologie také zahrnují různé hierarchické úrovně (viz např. položky „article“, „issue“ a „journal“). Přehled položek typologií je dostupný na jednom místě (vždy na jedné webové stránce). Obě typologie jsou stabilní, v dané podobě byly vytvořeny v roce 2004, od té doby nebyly aktualizovány.

Doplňme, že ve webovém rozhraní generátoru metadat ve formátu OpenURLCoins [251] je použito rozčlenění na celkem pět typů dokumentů, jimiž jsou „Journals and Articles“, „Books“, „Patents“, „Dissertations“ a „Generic“.

6.4.7 MARCXML

U formátu MARCXML si všimneme zejména typologie, která je uváděna v kódované podobě¹³⁵ v prvku „controlfield“ s hodnotou „006“ atributu „tag“. Vzhledem ke konstrukci formátu MARCXML (tj. východiskem je formát MARC 21) tyto hodnoty odpovídají hodnotám uváděným v poli 006 formátu MARC 21 [22].

Typologii tvoří celkem 15 položek, do značné míry jsou přitom navázané na obecné označení druhu dokumentu podle pravidel AACR2, o nichž je pojednáno v kap. 6.7.1. Obr. 41 názorně dokládá, že u jednotlivých položek jsou uplatňovány různé charakteristiky dokumentu (a jeho nosiče).



Obr. 41: Typologie dokumentů ve formátu MARCXML

V typologii jsou v čase prováděny určité změny, o nichž lze získat informace v části stránky s údaji o poli 006 formátu MARC 21 [22]. Typologie je

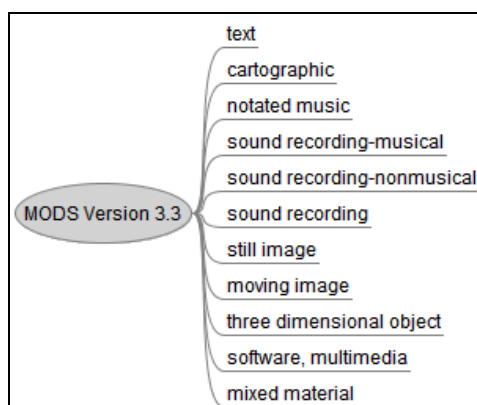
¹³⁵ S ohledem na srozumitelnost nejsou v tomto textu uváděny kódy příslušných položek, naopak je upřednostněno slovní vyjádření těchto kódů.

v zásadě stabilní, výše naznačené změny však ukazují, že v případě potřeby dochází k dílčím aktualizacím. Údaje o typologii a jednotlivých položkách, které ji tvoří, jsou k dispozici v rámci jedné webové stránky. Je však nutno konstatovat, že vysvětlivky týkající se použití jednotlivých položek jsou velmi stručné a mnohdy nepřinášejí uživateli podrobnosti o užití položky (např. vysvětlivka u položky „Notated music“ má podobu věty „Coded data elements relating to notated music.“) [22].

S položkami, jejichž kódované hodnoty jsou uváděny v poli 006 formátu MARC 21, resp. odpovídajícího prvku formátu MARCXML, jsou svázány kódované hodnoty v poli 008. Např. položkám „Language material“ a „Manuscript language material“ z pole 006 odpovídá položka „Books“ v poli 008.

6.4.8 MODS

Součástí formátu MODS je prvek „typeOfResource“, který může nabývat jedenácti hodnot (viz obr. 42) [252].



Obr. 42: Typologie dokumentů ve formátu MODS ve verzi 3.3

U jednotlivých typů lze spatřovat inspiraci v knihovnickém prostředí, konkrétně ve formátech typu MARC (viz kap. 6.4.7) a v pravidlech AACR2 (viz kap. 6.7.1).

U prvku „typeOfResource“ lze použít upřesňující atributy „collection“ a „manuscript“.

K typologii dokumentů má vztah ještě prvek „genre“, který může mít mj. atribut „authority“, v němž lze uvést hodnotu kódu žánru či formy z řízeného slovníku *Genre/Form Code and Term Source Codes* spravovaného Kongresovou knihovnou

ve Washingtonu. Jako hodnota prvku se pak uvádí konkrétní výraz z daného řízeného slovníku. Z určitého úhlu pohledu může v této souvislosti být zmíněn i prvek „physicalDescription“, resp. jemu podřazené prvky „form“ a „digitalOrigin“, popř. prvek „subject“ s podřazeným prvkem „genre“.

Formát MODS jako celek je průběžně aktualizován, tj. případné změny v typech dokumentů (v prvku „typeOfResource“) budou zachyceny v nových verzích formátu. Typologie je dostupná v XML schématu, přímo ve schématu je upřesněn způsob užití hodnoty „sound recording“. Typologii lze považovat za poměrně jednoznačnou, další podrobnosti o užití typů v praxi nabízejí pokyny pro uživatele [203].

6.4.9 NLM

Ve formátu NLM je k dispozici prvek „PublicationTypeList“, který může obsahovat jeden či více prvků „PublicationType“. Hodnoty, které je možno přidělit prvku „PublicationType“, jsou zároveň součástí hesláře MeSH¹³⁶, který spravuje NLM.

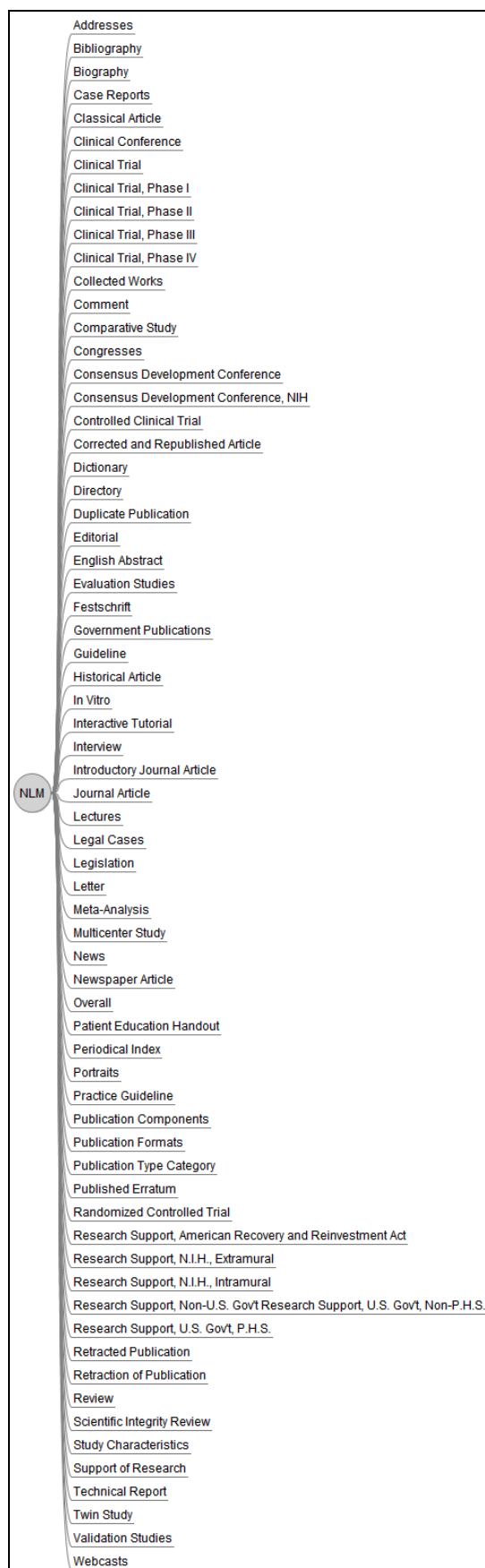
Jejich úplný výčet, který v současné době čítá 146 položek, je k dispozici v rámci tohoto hesláře [253] a zároveň v nápovědě k systému PubMed [254]. Na obr. 43 je vzhledem k vysokému počtu typů zachycen pouze jejich část, konkrétně se jedná o typy používané v systému PubMed [255].

Při zobrazení typů dokumentů v rozhraní hesláře MeSH (heslo „Publication Formats“) je k dispozici i vyjádření případných hierarchických vazeb mezi jednotlivými typy (na první úrovni je 85 typů, dalších 61 je pak na dalších úrovních).

Jako příklad můžeme uvést typ „Journal Article“, jemuž jsou na druhé úrovni podřazené typy „Introductory Journal Article“ a „Review“. Typ „Review“ je zároveň přímo podřazen prvku „Publication Formats“, neboť heslář MeSH využívá polyhierarchii. Na obr. 43 nejsou hierarchické vazby naznačeny, zmíněné typy jsou umístěny na stejné úrovni.

U jednotlivých typů dokumentů (stejně jako ostatních hesel MeSH) je k dispozici pole „Scope Note“, která poskytuje prostor pro upřesnění způsobu užití typu dokumentu (resp. obecně hesla MeSH).

¹³⁶ MeSH je zkratka názvu *Medical Subject Headings* (Lékařská předmětová hesla).



Obr. 43: Část typů dokumentů ve formátu NLM

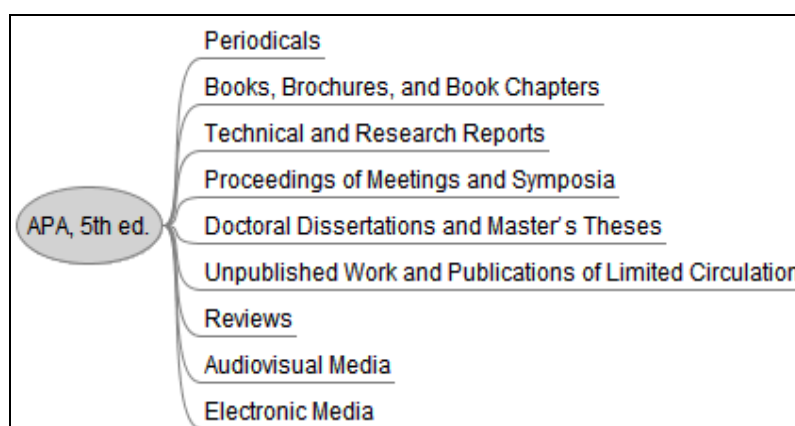
6.5 Typologie dokumentů v citačních stylech

6.5.1 APA

Pro naše účely se zaměříme na 5. a 6. vydání publikačního manuálu APA [25], [26], na tomto místě konkrétně na typologii dokumentů, s níž manuál pracuje.

Uživatelům 5. vydání manuálu je nabídnuta zjednodušená úvodní typologie, v níž se na nejvyšší úrovni setkáváme s pěti typy dokumentů („Periodicals“, „Nonperiodicals“, „Parts of a nonperiodical (e.g., book chapter)“, „Online periodicals“ a „Online documents“) ¹³⁷ [25, s. 223]. U periodik, neperiodik a online dokumentů jsou formou komentáře dále konkretizovány jejich podtypy.

Podrobnější typologie je uvedena spolu s konkrétními příklady záznamů [25, s. 239-281]. Je tvořena devíti typy dokumentů (viz obr. 44).



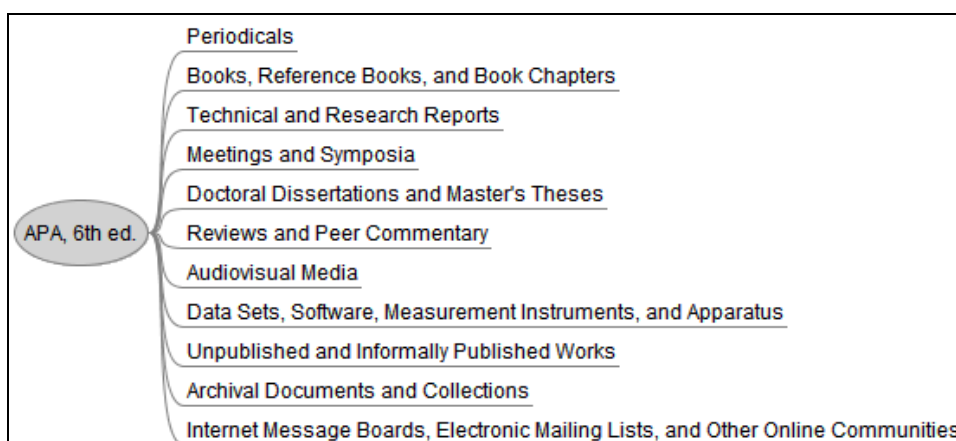
Obr. 44: Typologie dokumentů v publikačním manuálu APA (5th ed.)

Na další úrovni již nelze hovořit o typech (či subtypech) dokumentů, přesněji je označit položky na další úrovni jako popisky vystihující obecnější podstatu konkrétního uvedeného příkladu (v případě typu „Periodicals“ se jedná např. o „Journal article, two authors, journal paginated by issue“ či „Abstract from a secondary source (print periodical)“).

V 6. vydání manuálu [26, zejm. s. 193-224] se setkáváme s velmi podobnou typologií, která je ovšem složena již z jedenácti položek (viz obr. 45). Čtyři z nich

¹³⁷ Položku „Part of a nonperiodical (e.g., book chapter)“ lze pokládat za podřazenou položce „Nonperiodical“, v publikačním manuálu jsou však tyto položky uvedeny na stejné úrovni, proto je i na tomto místě zmíněno všech pět hlavních typů dokumentů.

jsou zcela shodné s typy dokumentů, s nimiž pracuje 5. vydání manuálu („Periodicals“, „Technical and Research Reports“, „Doctoral Dissertations and Master’s Theses“ a „Audiovisual Media“), další čtyři jsou velmi podobné („Books, Reference Books, and Book Chapters“, „Meetings and Symposia“, „Reviews and Peer Commentary“ a „Unpublished and Informally Published Works“). Významnější změna nastala v případě typu označeného jako „Internet Message Boards, Electronic Mailing Lists and Other Online Communities“, nově byly zařazeny položky „Data Sets, Software, Measurement Instruments, and Apparatus“ a „Archival Documents and Collections“. Podobně jako v 5. vydání manuálu jsou v manuálu uvedeny také položky na další úrovni mající charakter popisek, které vystihují obecnější podstatu konkrétního uvedeného příkladu



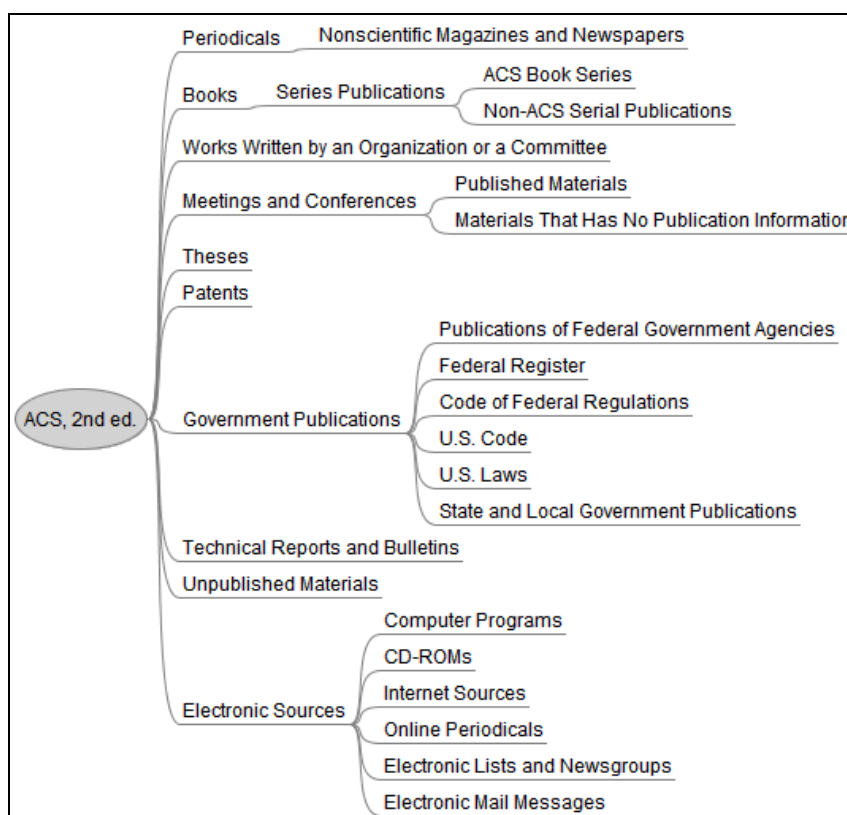
Obr. 45: Typologie dokumentů v publikačním manuálu APA (6th ed.)

Navíc jsou v 6. vydání doplněny vzory záznamů právnických dokumentů, ty však nejsou součástí stylu APA, ale jsou pouze převzaty z 18. vydání manuálu *Bluebook: A Uniform System of Citation*, který je využíván právníky. Proto nejsou v našem přehledu zařazeny.

6.5.2 ACS

Rovněž v případě publikačního manuálu (resp. průvodce stylem) ACS se budeme věnovat dvěma nejnovějším vydáním této příručky, tj. 2. a 3. vydání [27], [28].

Ve 2. vydání manuálu se v kapitole 6 „References“ setkáváme s typologií dokumentů, již tvoří celkem deset položek, z nichž u šesti jsou uvedeny i položky na další úrovni, u jedné i na třetí úrovni (viz obr. 46).

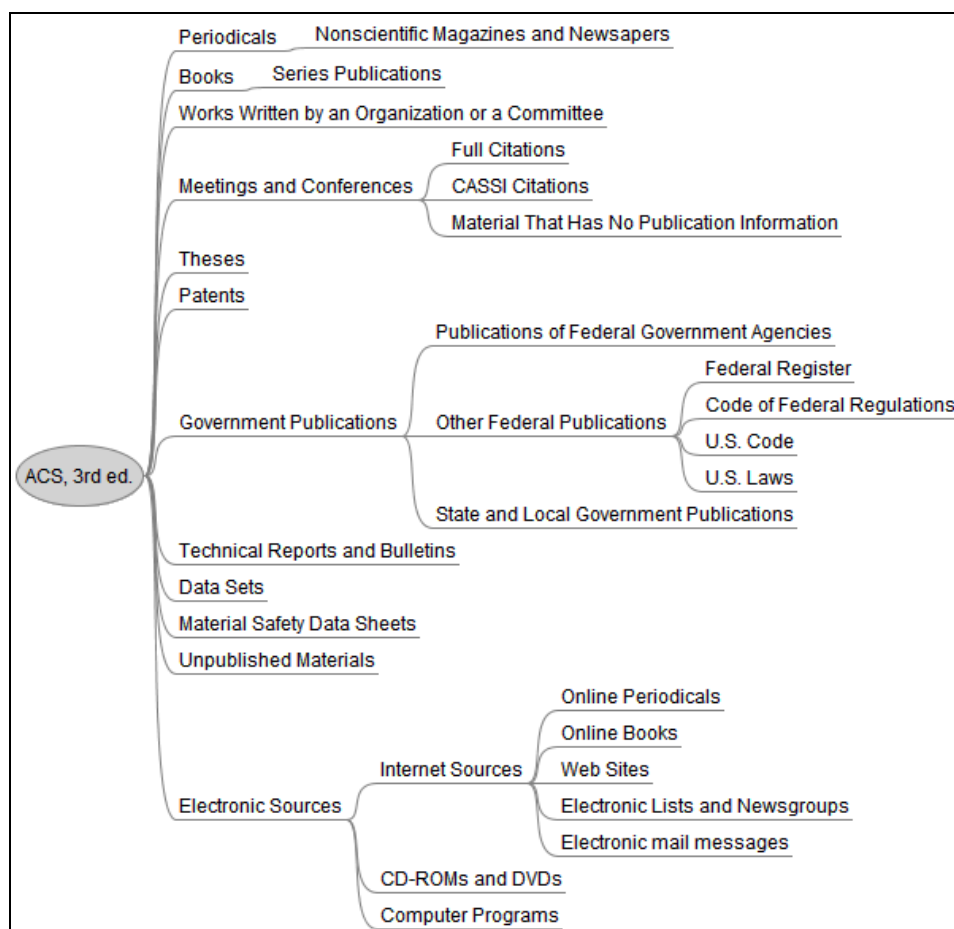


Obr. 46: Typologie dokumentů v publikačním manuálu ACS (2nd ed.)

Z obr. 46 je též zřejmé, že typologie je velmi pragmatická – je výrazně ovlivněna prostředím oboru, v němž vznikla (viz členění „Series Publications“), a rovněž prostředím ve smyslu geografickém (viz členění „Government Publications“).

Obdobnou podobu má i typologie dokumentů ve 3. vydání publikačního manuálu ACS (viz obr. 47) [28, s. 287-341]. Na první úrovni je složena z dvanácti položek, z toho se u čtyř setkáváme i s druhou hierarchickou úrovní členění a ve dvou případech i se třetí úrovní. Hlavními změnami je přidání dvou položek na nejvyšší úrovni („Data Sets“ a „Material Safety Data Sheets“), vyzdvižení dvou typů internetových zdrojů jako samostatných položek („Online Books“ a „Web Sites“), přidání výrazu „and DVDs“ k původní položce „CD-ROMs“ a zpřesnění hierarchie

(jedná se o přidání položky „Other Federal Publications“ a o podřazení jednotlivých typů internetových zdrojů typu „Internet Sources“).



Obr. 47: Typologie dokumentů v publikačním manuálu ACS (3rd ed.)

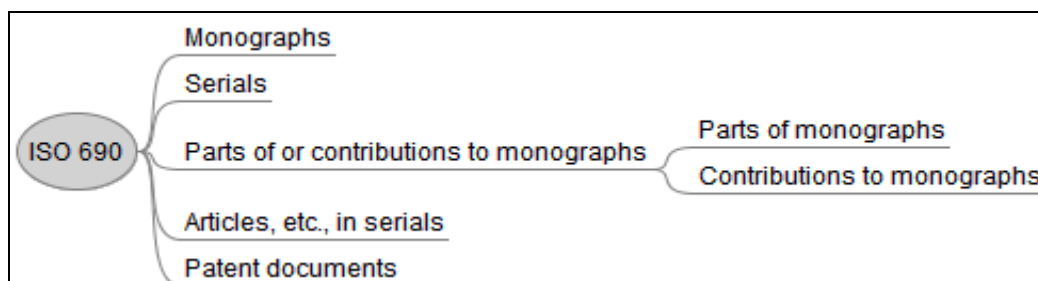
Poznamenejme, že přehledy typů byly extrahovány přímo z textu, resp. uvádění jednotlivých úrovní nadpisů (vzájemně typograficky odlišených). Ve 3. vydání je (na rozdíl od 2. vydání) totiž ještě k dispozici tabulka s obvyklými typy záznamů s příklady, kde jsou některé typy vymezeny jinak (resp. upřesněny podle případů vyskytujících se typicky v praxi) [28, s. 292-293].

6.5.3 ISO 690 a 690-2

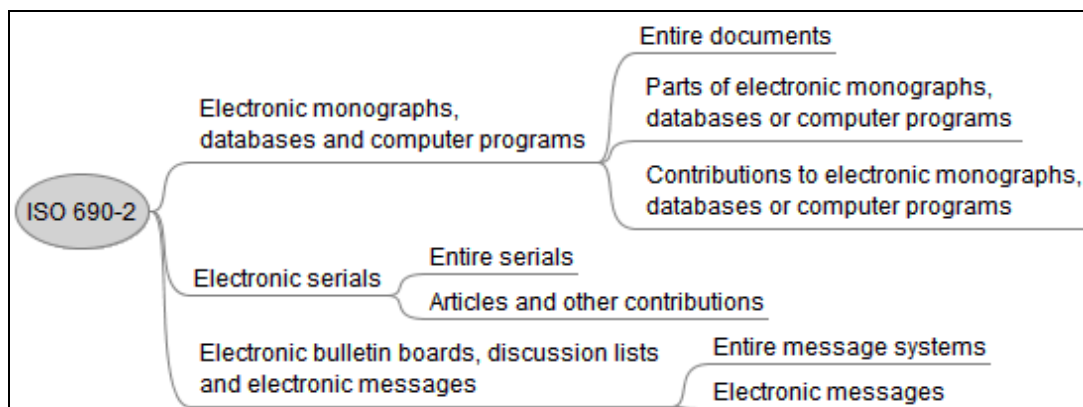
V mezinárodních normách ISO 690 [29] a ISO 690-2 [30] (resp. také v jejich českém překladu¹³⁸ [52], [53]) je uvedena velmi jednoduchá typologie

¹³⁸ Z důvodu jazykové konzistence typologií charakterizovaných v této práci je i typologie převzatá z těchto norem uvedena v anglickém jazyce. České překlady obou norem používají stejnou typologii, proto v tomto konkrétním případě není bezpodmínečně nutné rozlišovat mezi původními anglickými verzemi norem a jejich českými překlady.

dokumentů (viz obr. 48 a 49) – v případě první zmíněné normy se na první úrovni jedná o pět typů dokumentů (pouze v jednom případě se setkáme s rozšířením na další úroveň), v případě druhé normy o tři typy dokumentů (z nich ovšem všechny jsou podrobněji rozčleněny na další úrovni).¹³⁹



Obr. 48: Typologie dokumentů v normě ISO 690



Obr. 49: Typologie dokumentů v normě ISO 690-2

Typologie jsou dostupné v rámci textů norem, tyto primární texty však nejsou k dispozici zdarma prostřednictvím internetu. Jsou však k dispozici nejrůznější příručky s interpretacemi norem a s praktickými ukázkami jejich využití.

Typologii lze pokládat za relativně jednoznačnou, u jednotlivých typů dokumentů jsou uvedeny vysvětlivky a příklady. Jako problematičtější se jeví zejména přiřazování nových typů dokumentů k typům obsaženým v normách (na první pohled není např. zřejmé, ke kterému šířeji pojatému dokumentu přiřadit příspěvek na blogu) a zcela jiných typů dokumentů (např. norem). V praxi je proto nutné, aby uživatel citující podle těchto norem využil příručky, v nichž jsou

¹³⁹ Označení typů dokumentů je sjednoceno tak, že se vždy jednalo o plurál – plurálové označení typů v normách převládá, singulár se vyskytuje pouze ojediněle.

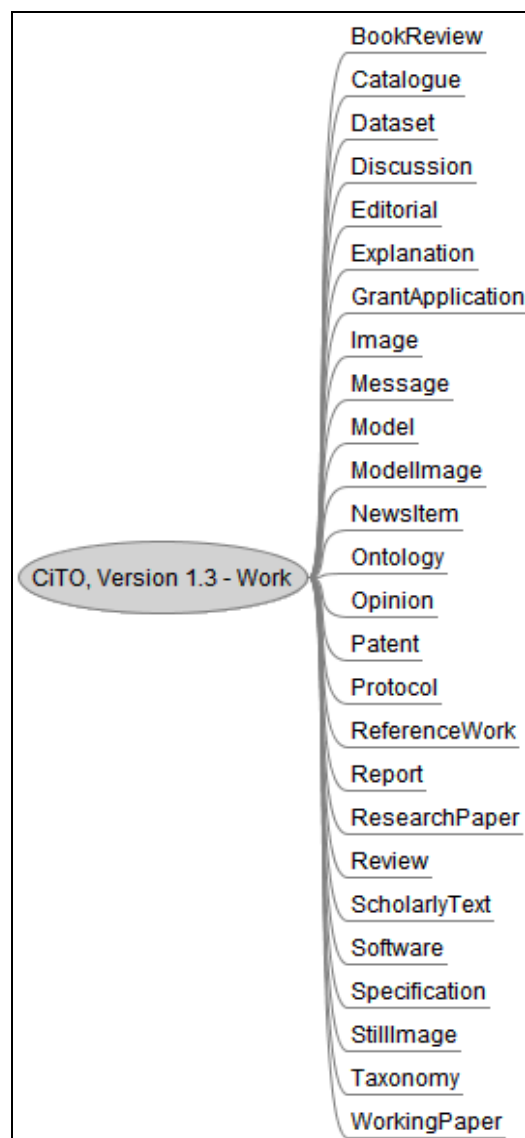
doporučené způsoby citování některých typů dokumentů explicitně do normy nezahrnutých) nebo při přiřazování uplatnil vlastní úsudek.

Aktualizace těchto typologií je realizována formou nové verze normy ISO 690 (*Information and documentation – Guidelines for bibliographic references and citations to information resources*), která byla vydána 14. června 2010 a aktualizuje obě výše zmíněné normy. (Vzhledem k datu vydání normy nebylo již možné do předkládané práce aktualizovanou typologii zahrnout.)

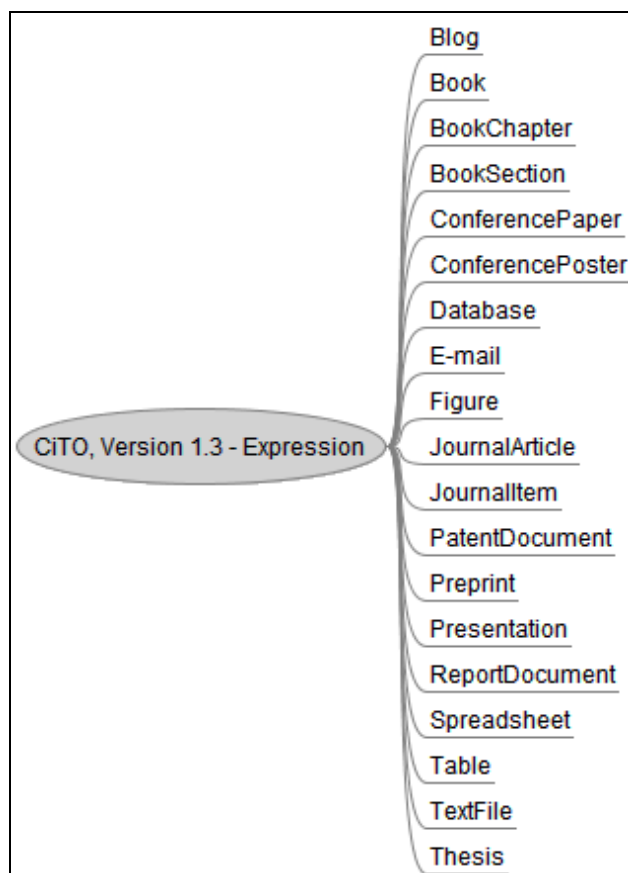
6.6 Typologie dokumentů v citačních ontologiích

6.6.1 CiTO

Typologie dokumentů v ontologii CiTO myšlenkově využívá FRBR. Ve verzi 1.3 ontologie se na první úrovni typologie jedná o celkem 26 typů díla (*Work*) (obr. 50), na druhé úrovni o celkem 19 typů vyjádření (*Expression*) (obr. 51) a na třetí úrovni o čtyři typy provedení (*Manifestation*) (obr. 52) [31].

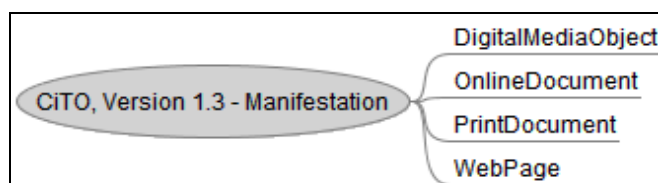


Obr. 50: Typologie dokumentů v ontologii CiTO
– úroveň Work (dílo)



Obr. 51: Typologie dokumentů v ontologii CiTO – úroveň Expression (vyjádření)

Bude-li se tedy např. jednat o tištěný časopisecký článek ve vědeckém časopise, je možné uvažovat o jeho zařazení jako díla „ResearchPaper“, vyjádření „JournalArticle“ a provedení „PrintDocument“. Ontologie CiTO navíc umožňuje zachycení statusu dokumentu z hlediska recenzování (možnosti *peerReviewed* a *notPeerReviewed*). Díky tomu lze typologii pokládat za relativně jednoznačnou.



Obr. 52: Typologie dokumentů v ontologii CiTO – úroveň Manifestation (provedení)

6.6.2 BIBO

Specifikace ontologie BIBO [107], [6] uvádí celkem 69 tříd¹⁴⁰ – z hlediska typologie dokumentů je důležitých 57 z nich (jedná se o třídy s prefixem bibo)¹⁴¹, viz obr. 53.

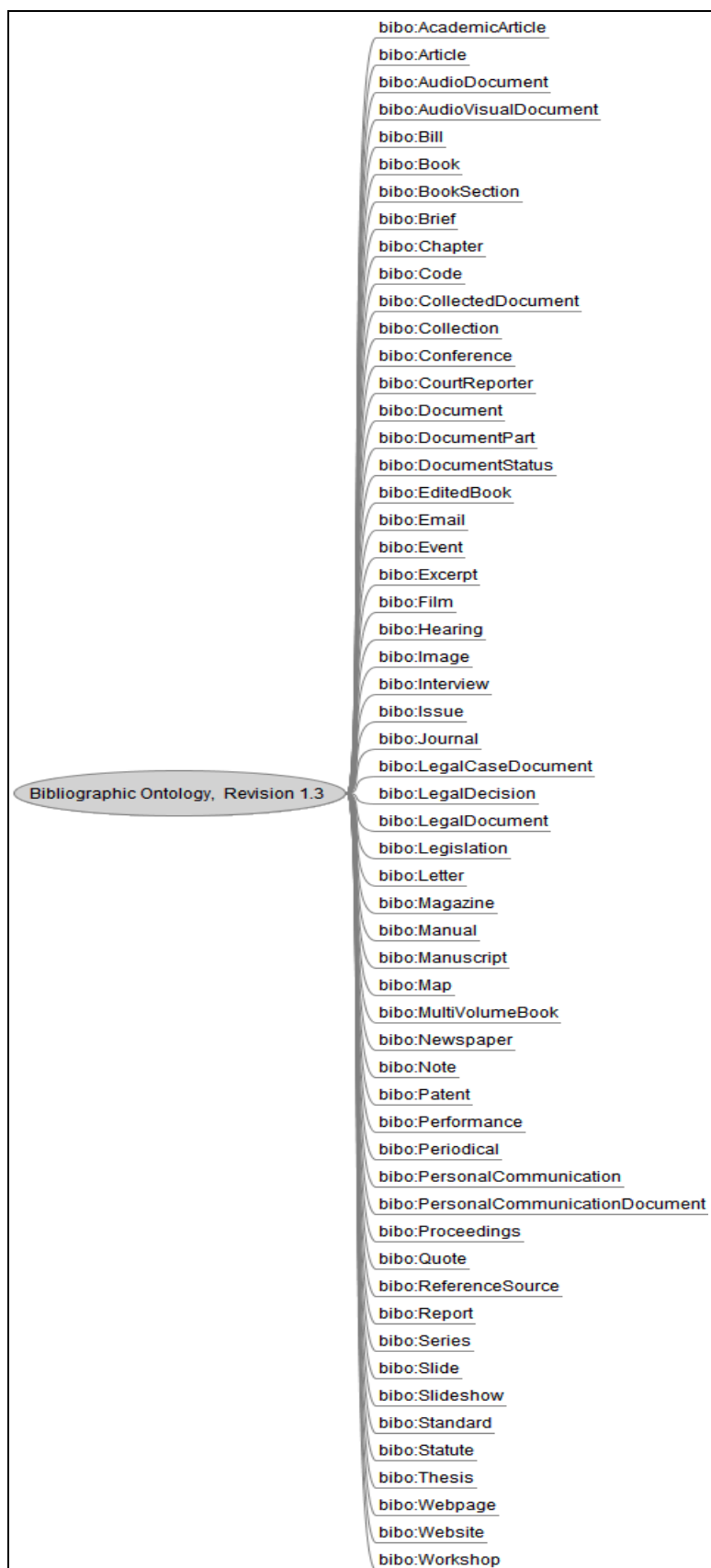
Významným pozitivem typologie, kterou přináší tato ontologie, je podrobný komentář týkající se vhodného užití jednotlivých typů, resp. tříd. Jako příklad je možné zmínit komentáře náležející ke třídám „bibo:AcademicArticle“ a „bibo:Article“ – zatímco v prvním případě se jedná o vědecký článek typicky zveřejněný v časopise, v případě druhém je očekáváno použití v případě, kdy se jedná o převážně neumělecký text na konkrétní téma, který tvoří nezávislou část knihy nebo jiné publikace, např. novin či magazínu.

V ontologii BIBO lze také v určitých případech vyjádřit vztah celek-část, kdy např. číslo časopisu je součástí časopisu jako celku.

Podobně jako v případě ontologie CiTO je i ontologie BIBO jako celek je podrobována aktualizacím, takže i typy dokumentů mohou být průběžně aktualizovány. Aktuální typologie je přitom dostupná v rámci aktuální verze celé ontologie, tj. v jednom souboru.

¹⁴⁰ Součástí ontologie BIBO nejsou však pouze třídy, ale také vlastnosti objektů (*Object Properties*), vlastnosti dat (*Data Properties*) a jednotlivosti (*Individuals*).

¹⁴¹ Do přehledu nejsou zahrnuty třídy „owl:Thing“ (kořenová třída), dále „bibo:ThesisDegree“ (jediná nezařazená třída s prefixem bibo) a třídy s prefixy dcterms, event, foaf, rdf a rdfs.

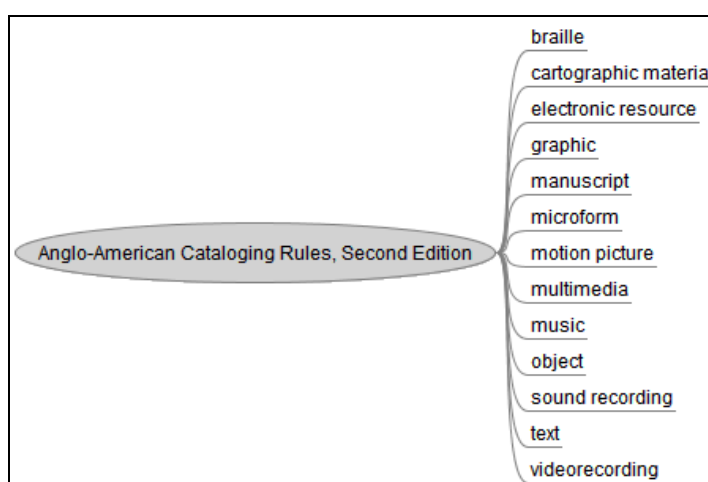


Obr. 53: Typologie dokumentů v ontologii BIBO

6.7 Typologie dokumentů v oblasti knihovnictví

6.7.1 AACR2

Velmi pragmatickou typologií, v níž se mísí několik hledisek, z nichž lze pohlížet na dokument, a v níž je spor obsah versus nosič zvláště patrný, nabízí v současné době platné 2. revidované vydání AACR, tj. AACR2 [256].¹⁴² V pravidle 1.1C1 jsou uvedeny dva seznamy obecného označení druhu dokumentu¹⁴³, z nichž první obsahuje třináct položek a druhý 27 položek. V ČR je využíván seznam č. 1 [257] (viz obr. 54).



Obr. 54: Jedna z typologií dokumentů v Anglo-amerických katalogizačních pravidlech

Přestože je v pravidlech užíván výraz druh dokumentu, domnívám se, že v tomto případě je příhodnější výraz typ dokumentu, neboť na dokument je pohlíženo z různých hledisek – v dalším textu je proto již upřednostňován výraz typ dokumentu.

Stejně jako další části těchto pravidel podléhá i uvedená typologie dokumentů změnám v čase – pravidla jsou aktualizována formou vydávání dodatků (*amendments*). Z hlediska typologie dokumentů je významný dodatek k pravidlům

¹⁴² Určitou typologii dokumentů (z hlediska potřeb katalogizace) nastiňuje (zejména svým rozdělením na kapitoly) i původní vydání AACR (viz zejména „Chapter 6. Separately Published Monographs“, „Chapter 7. Serials“, „Chapter 8. Incunabula“, „Chapter 9. Photographic and Other Reproductions“, „Chapter 10. Manuscripts“, „Chapter 11. Maps, Atlases, etc.“, „Chapter 12. Motion Pictures and Filmstrips“, „Chapter 13. Music“ a „Chapter 14. Phonorecords“, „Chapter 15. Pictures, Designs, and Other Two-Dimensional Representations“) [258].

¹⁴³ Pro obecné označení druhu dokumentu je používána zkratka GMD (*General Material Designation*).

z roku 2001, který mj. stanovil změnu označení typu počítačový soubor (*computer file*) na elektronický zdroj (*electronic resource*) [32], [259]. Podle B. Tillettové se jednalo spíše o kosmetickou změnu, kterou však lze považovat za užitečnou pro katalogizátory i uživatele [260, s. 169].

V případě, že popisná jednotka obsahuje více než jeden typ dokumentu a ani jeden z nich nepřevažuje, nebo pokud se jedná o soubor textových dokumentů na jednom médiu, má být podle poznámky k pravidlu 1.1C1 a podle pravidla 1.1C4 uvedeno označení „multimedia“ nebo „kit“).

S ohledem na skutečnost, že pravidla jsou aktualizována formou dodatků, lze říci, že není vždy snadné získat ověřenou aktuální podobu typologie – dodatky k pravidlům jsou vydávány formou volných listů, jimiž je v tištěné verzi pravidel třeba postupně nahrazovat původní znění pravidel. Existuje sice také online verze pravidel, ta je však dostupná pouze komerčně, a to prostřednictvím online služby Cataloger's Desktop [261]. Stručná shrnutí obsahu jednotlivých dodatků jsou k dispozici na webu Společného řídicího výboru pro vývoj RDA (*Joint Steering Committee for Development of RDA*)¹⁴⁴ [262].

S ohledem na připravovanou výraznou změnu katalogizačních pravidel, která se promítla i do změny názvu [263] lze usuzovat, že typologie nebude dále rozvíjena.

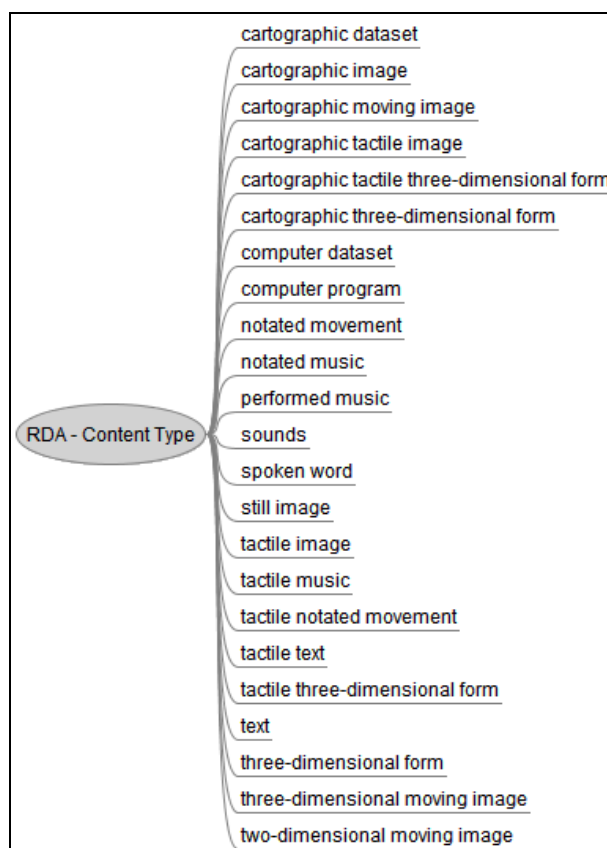
6.7.2 RDA

V návrhu pravidel RDA [33], které vycházejí z principů FRBR [112], se s problematikou typologie dokumentů setkáváme na několika místech textu, a to z několika úhlů pohledu. Podstatná je zejména kapitola 3 „Describing Carriers“ a kapitola 6.10 „Content Type“. V kapitole 3 je na dokumenty nahlíženo z pohledu typu (*type*), rozsahu (*extent*) a rozměrů (*dimensions*) nosiče, zvláštní pozornost je věnována i typu média (*media type*). Pokud se týká online zdrojů, nejsou dále výrazněji členěny, je pouze předepsáno přiřadit jim typ „online resource“ jako typ nosiče.

V kapitole 6.10 je pak doporučen způsob zachycení typu obsahu (*content type*). Ten v podstatě odpovídá obecnému označení druhu dokumentu, s nímž pracují

¹⁴⁴ RDA je zkratka výrazu *Resource Description and Access* (Popis a zpřístupnění zdrojů).

dosud platná pravidla AACR2. Je-li typologie využívána pro účely katalogizace (a v souladu s nastíněnými pravidly), je možné buď uvést všechny typy aplikovatelné na daný dokument (nebo jeho podstatné části), nebo typ, který odpovídá nejdůležitější části zdroje. Typologie je složena z 23 položek (viz obr. 55). Navíc je ještě v případech, kdy není vhodná žádná z nabízených položek, resp. nemůže být zjištěna, možné použít položku „other“, resp. položku „unspecified“.



Obr. 55: Typy obsahu v pravidlech RDA

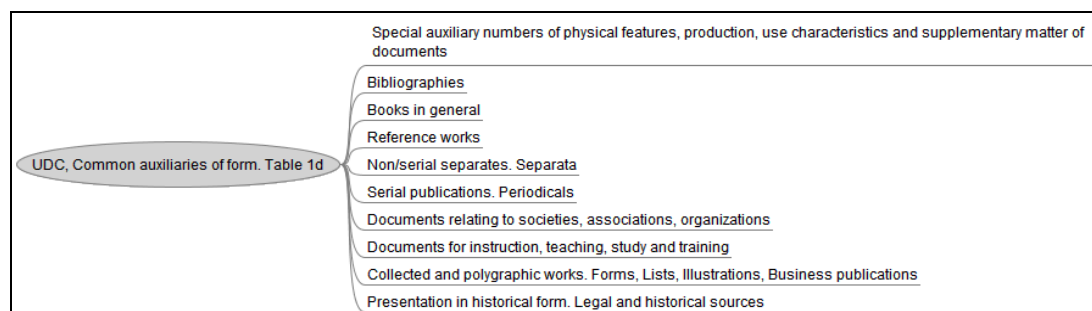
Pokud se týká způsobu aktualizace typologie, lze očekávat, že aktualizace bude součástí aktualizace celých pravidel RDA. Aktuální verze pravidel však bude po svém schválení zpoplatněna [264].

Omezíme-li se pro naše účely na typy obsahu, lze říci, že potřebné údaje jsou k dispozici na jednom místě v rámci pravidel. Budeme-li však uvažovat i o typech vztahujících se k nosičům, resp. médiím, pak je nutno konstatovat, že údaje jsou uvedeny v rámci dvou kapitol pravidel.

Jako pozitivní lze označit skutečnost, že k jednotlivým typům obsahu je vždy k dispozici komentář upřesňující očekávaný způsob užití.¹⁴⁵

6.7.3 MDT

S velmi podrobnou typologií dokumentů se setkáme také v tabulce Id samostatných všeobecných pomocných znaků formy dokumentu v MDT¹⁴⁶. Ty jsou určeny pouze k vyjadřování formy dokumentu, která se týká vnějšího fyzického vzhledu, písma a grafického (typografického) uspořádání a způsobu zpracování, nikoliv jeho obsahu (předmětu). Lze-li u dokumentu identifikovat více forem, je rozlišováno mezi vnitřní a vnější formou [267, s. 21]. Přehled slovních vyjádření znaků na nejvyšších úrovních z tabulky 1d je k dispozici na obr. 56.



Obr. 56: Typologie znaků formy v MDT

Obecně lze konstatovat, že tabulky MDT se průběžně aktualizují, v současné době pod patronátem Konsorcia MDT¹⁴⁷, dříve pod hlavičkou FID¹⁴⁸, která navázala na činnosti Mezinárodního bibliografického ústavu¹⁴⁹,

¹⁴⁵ V souvislosti s RDA a typologií dokumentů je vhodné zmínit také vytvoření společného rámce pro kategorizaci zdrojů RDA/ONIX [265]. ONIX (*ONline Information eXchange*, Výměna online informací) je XML formát využívaný vydavateli pro distribuci metadat [266].

¹⁴⁶ MDT je zkratka názvu Mezinárodní desetinné třídění. V angličtině je pro toto třídění využívána zkratka UDC (*Universal Decimal Classification*).

¹⁴⁷ Konsorcium MDT (*UDC Consortium*) sídlí v Královské knihovně v Haagu v Nizozemsku (*Koninklijke Bibliotheek*).

¹⁴⁸ FID je zkratka názvu *Fédération Internationale de Documentation* (Mezinárodní federace pro dokumentaci).

¹⁴⁹ Mezinárodní bibliografický ústav (*Institut International de Bibliographie*) byl několikrát přejmenován. Nejvýznamnější změna nastala ve druhé polovině 30. let 20. století, kdy byla organizace transformována na Mezinárodní federaci pro dokumentaci. Na konci 80. let byl do názvu přidán informační aspekt (ve francouzštině název zněl *Fédération Internationale d'Information et Documentation*). Organizace existovala do roku 2002, kdy byla rozpuštěna, resp. stala se součástí Mezinárodní federace knihovnických sdružení a institucí (*International Federation of Library Associations and Institutions*, IFLA). Správa a rozvoj MDT však přešly pod patronát Konsorcia MDT již v roce 1992 [269], [270, s. 290-294], [271], [272].

v němž byla původní verze MDT (vycházející z Deweyho desetinného třídění) vytvořena¹⁵⁰.

Konkrétně jsou návrhy na změny tabulek a nejnovější schválené změny každoročně publikovány pod názvem *Extensions and corrections to the UDC* [273]. Navíc je přehled všech zrušených položek publikován na webu Konsorcia MDT [274], [275].

Konsorciem udržuje tzv. Hlavní referenční soubor¹⁵¹, který je jednou ročně aktualizován a jsou do něj zahrnovány schválené změny. Soubor představuje autorizovanou verzi MDT a je dostupný na základě licence [274].

Zároveň je třeba konstatovat, že vzhledem k rozsahu úplného třídění jsou v různých jazycích vydávány různé verze třídění – od úplného přes střední až po zkrácené. Např. výše citované české vydání MDT [267] je příkladem zkráceného vydání.

Ve třídě (0.0...) Zvláštní pomocné znaky se setkáme s členěním typů dokumentů podle různých kritérií (podle fyzické, vnější formy, podle způsobu výroby, podle stadia výroby, pro jednotlivé druhy uživatelů, podle úrovně podání a dostupnosti).

Typologii dokumentů, s níž prostřednictvím tabulky samostatných všeobecných pomocných znaků pracuje MDT, lze pokládat za postupně rozšiřovanou.¹⁵² V úplnosti je typologie dostupná v rámci nově vznikajícího systému UDC Summary [34] dostupného pod licencí Creative Commons Attribution Share-Alike 3.0 Unported [276], popř. v rámci online vydání českého překladu MDT na serveru Národní knihovny ČR [277].

6.8 Dílčí závěry

Typologie dokumentů představuje jedno z často diskutovaných témat v odborné literatuře – zabývali se jím již klasikové jako P. Otlet nebo B. C. Vickery. Obsah a rozsah pojmu označovaného termínem „dokument“ či příbuznými termíny

¹⁵⁰ První úplné vydání MDT pochází z roku 1905, proto je tento rok pokládán za rok skutečného vzniku MDT [268, s. 74].

¹⁵¹ Je označován jako *Master Reference File* (MRF).

¹⁵² Názorným příkladem může být položka (0.034) Strojem čitelné dokumenty. S řadou změn týkajících se mj. i znaků formy dokumentu se lze seznámit v rámci přehledu zrušených položek [275].

typu „informační objekt“ se rovněž stal tématem řady odborných textů z oblasti informační vědy. Zvláště s rozvojem elektronického publikování a vznikem řady nových typů dokumentů se těmito úvahami zabývá řada autorů, a to jak v zahraničním, tak v domácím prostředí.

V této práci je za dokument považován informační pramen, který existuje v hmotné nebo nehmotné podobě, obsahuje množinu informací a slouží k přenosu těchto informací v čase a prostoru.

Lze shrnout, že v oblasti zkoumaných citačních formátů, stylů, ontologiích ani ve vybraných zdrojích užívaných v knihovnické praxi nejsou používány stejné typologie dokumentů, resp. liší se podle primárního cíle, pro který byly vytvořeny či pro který jsou určeny. V tomto směru lze výše rozebírané typologie dokumentů pokládat za spíše pragmatické či účelové.

Pro účely formulace doporučených postupů pro zahrnutí citačních dat do XML formátů a návaznou tvorbu experimentálního XML formátu pro zachycení těchto dat se proto jako optimální řešení jeví převzetí některé z existujících typologií nebo vytvoření nové typologie a její namapování na další již existující typologie.

Za vhodné lze pokládat typologie dokumentů, které jsou volně dostupné, nejlépe na jednom místě na internetu, postupně rozšiřované, doplněné o komentář týkající se použití typů dokumentů v praxi.

Z výše charakterizovaných typologií se v tomto směru jako zvláště příhodné jeví typologie citačních formátů MODS a NLM, ontologií BIBO a CiTO, popř. i třídění MDT. Problematictější naopak jsou typologie citačních formátů bez vysvětlivek týkající se použití formátu v praxi nebo pouze s nepřímými vysvětlivkami (RIS, RefWorks, EndNote a MARCXML) či bez oficiálních aktualizací (BibTeX). Za optimální typologie dokumentů dále nelze považovat typologie v citačních stylech, které jsou sice v nových vydáních aktualizovány, avšak jsou typicky rozptýleny na mnoha stranách manuálů a není vždy zřejmá hierarchie jejich prvků (APA a ACS). Typologie v normách ISO 690 a 690-2 je pak problematická zejména s ohledem na svou obtížnou dostupnost. Podobně je tomu i s typologií dokumentů v pravidlech AACR2 a v připravované oficiální verzi pravidel RDA.

7. Citační praxe v tzv. barevných knihách IUPAC

7.1 Úvod

V této kapitole jsou prezentovány výsledky analýzy citační praxe uplatňované v tzv. barevných knihách IUPAC, konkrétně v jejich online verzích dostupných prostřednictvím internetu.

Analýza současného stavu online barevných knih z hlediska citování byla provedena pro potřeby výboru CPEP [278]. Představuje jeden z klíčových podkladů pro návrh XML formátu pro zachycení citačních údajů v barevných knihách. Tento formát bude využit v publikačním systému organizace IUPAC, resp. v informačním systému, jehož hlavním výstupem z hlediska uživatele bude webová strojově srozumitelná prezentace barevných knih připravená s využitím technologií založených na jazyce XML. Do strojově srozumitelné podoby je zatím zpracována Zlatá kniha, je připravován převod Červené a následně Zelené knihy [279].¹⁵³

Nejprve je stručně představena organizace IUPAC, a to zejména s ohledem na své publikační aktivity, dále je již práce zaměřena na zmíněné barevné knihy – jsou prezentovány výsledky hodnocení online verzí těchto knih z hlediska citační praxe. V dílčím závěru jsou stručně shrnuty nejdůležitější výsledky provedené analýzy z pohledu dalších plánovaných prací týkajících se strojově srozumitelných verzí barevných knih.

7.2 IUPAC a publikační činnost

IUPAC je organizace, jejímž posláním je rozvíjet celosvětové aspekty chemických věd a přispět k použití chemie ve službách lidstva. Vznikla v roce 1919, v roce 2009 tedy oslavila již 90. výročí svého vzniku. Její význam spočívá ve sjednocování „jazyka“ chemiků – hraje vůdčí roli např. v oblastech chemické nomenklatury, terminologie, standardizovaných metod měření či atomových vah [280].

¹⁵³ Pro tyto účely je plánováno využití formátu EPUB, tj. otevřeného standardu pro elektronické knihy [143].

Organizace IUPAC produkuje řadu různých typů dokumentů – jejich přehled (resp. rozcestník k nim) je k dispozici na webu organizace [281] v sekci „Publications“ (viz obr. 57).

IUPAC publications

Books and reports

[Technical Reports and Recommendations](#)
reports and recommendations endorsed by IUPAC

[IUPAC Books](#)
including nomenclature books

Journals

[Chemistry International](#)
The news magazine of IUPAC is published bimonthly to keep you informed about IUPAC's activities.

[Pure and Applied Chemistry](#)
The IUPAC's official journal is published monthly and contains recommendations, reports, and lectures from symposia.

[Macromolecular Symposia](#)
Contributions in the field of macromolecular chemistry and physics from selected international meetings, including those sponsored by IUPAC.

[Solubility Data Series](#)
This series is published in issues of the Journal of Physical and Chemical Reference Data (JPCRD).

E-resources

[Nomenclature and terminology](#)
on-line access to IUPAC nomenclature and terminology books

[Databases](#)
solubilities, activity coefficients, kinetic data, stability constants, and more

[Educational resources](#)
presentations, collections of diagrams, and more

[InChI](#)
open source software for automatic generation of IUPAC International Chemical Identifiers

[ThermoML](#)
IUPAC sponsored XML language for thermodynamic property data storage.

[Chemical Education International](#)
The online newsletter of the Committee on Chemistry Education presents reports, facts and figures, as well as opinions and essays.

Obr. 57: Rozcestník k publikacím organizace IUPAC

Přestože jsou dokumenty rozčleněny do tří hlavních skupin podle svého typu na Knihy a zprávy (*Books and Reports*), Časopisy (*Journals*) a E-zdroje (*E-resources*), je nutné upřesnit, že obsah jednotlivých skupin se mnohdy prolíná. Jako příklad lze uvést podskupinu Technické zprávy a doporučení skupiny Knihy a zprávy – zprávy a doporučení jsou převážně zároveň publikovány v časopise *Pure and Applied Chemistry*, který je pochopitelně zařazen ve skupině Časopisy. Uvedený přehled publikovaných dokumentů na jednu stranu umožňuje uživatelům

vícekriteriální přístup k dokumentům, na druhou stranu skutečnost, že jeden dokument lze nalézt různými způsoby, může působit matoucím dojmem.

Pro naše účely se zaměříme na tzv. barevné knihy (*color books*). Podotkněme, že s ohledem na orientaci práce na strojově zpracovatelná data se budeme zabývat primárně elektronickými verzemi publikačních výstupů.

Tzv. barevné knihy obsahují chemickou nomenklaturu. Jednotlivé barvy přitom odlišují zaměření titulů [282]:

- Zlatá kniha (*Gold Book*)¹⁵⁴ – Chemická terminologie (*Chemical Terminology*);
- Zelená kniha (*Green Book*) – Množství, jednotky a symboly ve fyzikální chemii (*Quantities, Units and Symbols in Physical Chemistry*);
- Červená kniha (*Red Book*) – Nomenklatura anorganické chemie (*Nomenclature of Inorganic Chemistry*);
- Modrá kniha (*Blue Book*) – Nomenklatura organických sloučenin (*Nomenclature of Organic Compounds*);
- Fialová kniha (*Purple Book*) – Makromolekulární nomenklatura (*Macromolecular Nomenclature*);
- Oranžová kniha (*Orange Book*) – Analytická nomenklatura (*Analytical Nomenclature*);
- Stříbrná kniha (*Silver Book*) – Nomenklatura a symboly v klinické chemii (*Nomenclature and Symbols in Clinical Chemistry*);
- Bílá kniha (*White Book*) – Biochemická nomenklatura a související dokumenty (*Biochemical Nomenclature and Related Documents*).

Jak již bylo naznačeno, v následujících částech textu se budeme těmito publikacemi zabývat podrobněji, a to z hlediska uplatnění citačních principů. Konkrétně se jedná o dvě základní oblasti, resp. dva případy:

- v knize jsou citovány jiné dokumenty, přičemž podrobné bibliografické údaje o nich jsou typicky uváděny v závěru dané publikace a v textu je na tyto podrobné záznamy odkazováno;

¹⁵⁴ Zároveň se jedná o jméno – Victor Gold se zasloužil o první vydání této knihy (GOLD, V.; LOENING, K. L. *Compendium of Chemical Terminology : IUPAC Recommendations*. Oxford : Blackwell, 1987. 456 s.) – tj. možný je i překlad *Goldova kniha*.

- kniha sama je předmětem citování.

7.3 Barevné knihy

Přehled všech barevných knih s důrazem na jejich online verze je k dispozici na webu organizace IUPAC [283]. Řada z těchto barevných knih je k dispozici ve více jazycích, v následujícím textu se zaměříme na jejich původní anglické verze.

Text týkající se jednotlivých barevných knih je vždy (s výjimkou Stříbrné knihy) rozdělen do dvou podkapitol, jimiž jsou Záznamy použité literatury a Bibliografické údaje o knize.

První podkapitola je zaměřena na způsob práce s citacemi v textu a se záznamy použitých zdrojů. Je-li to možné, jsou uvedeny počty záznamů.¹⁵⁵

Protože v případě barevných knih nelze (na rozdíl od časopisů vydávaných organizací IUPAC¹⁵⁶) vycházet z instrukcí pro autory, jsou v prvních podkapitolách obsaženy pouze ukázky záznamů použitých zdrojů a jejich citací v textu. S ohledem na možnost vzájemného porovnání záznamů byl v tomto případě zvolen jeden konkrétní typ dokumentu, a to časopisecký článek. Ukázky záznamů časopiseckých článků jsou doplněny o komentáře týkající se uvedených údajů a jejich formální podoby a jsou uváděny v textu; v případě, že se jejich podoba v jednotlivých částech dané knihy liší, jsou jejich příklady uváděny v tabulce.

Ve druhých podkapitolách jsou rozebrány způsoby uvádění bibliografických údajů o knize, a to v podobě přímo viditelné běžnému uživateli i v podobě skryté.

7.4 Zlatá kniha – Chemická terminologie

Zlatá kniha je zatím jako jediná z barevných knih IUPAC zpracována ve formátu XML, tj. v pokročilé strojově srozumitelné verzi [159]. Pro její zpracování byl vyvinut program Goldify [288]. K dispozici je rovněž informační stránka o Zlaté knize [289].

¹⁵⁵ Pro další fázi přípravy strojově srozumitelných verzí barevných knih hrají orientační úlohu a při vlastní práci bude jejich počet zpřesněn.

¹⁵⁶ Nejdůležitější roli hraje časopis *Pure and Applied Chemistry* [284] (z časopisů vydávaných organizací IUPAC se u tohoto titulu v pokynech pro autory setkáváme s nejpřesnějšími instrukcemi z hlediska citování [285]); dále organizace vydává zpravodaje *Chemistry International* [286] a *Chemical Education International* [287].

7.4.1 Záznamy použité literatury a citace

Přehled záznamů zdrojových dokumentů je k dispozici v sekci nazvané „source documents“ [290]. Záznamy jsou rozděleny do čtyř skupin:

- Knihy (*Books*);
- ISO (*ISO*);
- Pure and Applied Chemistry (časopis vydávaný organizací IUPAC);
- Ostatní (*Other*).

U jednotlivých záznamů článků z časopisu Pure and Applied Chemistry jsou v hranatých závorkách zároveň uváděny počty citací ve Zlaté knize; rovněž v hranatých závorkách je u těchto záznamů doplněn příslušný odkaz na online verzi dokumentu (směřuje z výrazu „Full text“). U posledního záznamu v první skupině (tj. Knihy) je rovněž uvedeno URL, obsahuje jej výraz označující vydavatele (National Institute of Standards and Technology).

Ve skupině Knihy je obsaženo celkem devět záznamů, ve skupině ISO pět záznamů, ve skupině Pure and Applied Chemistry 131 záznam a ve skupině Ostatní dva záznamy. Celkem se tedy jedná o 147 záznamů.

Následující příklad citovaného časopiseckého článku ukazuje, že nejprve je uvedena zkratka názvu časopisu „PAC“, dále následují rok vydání, ročník a číslo stránky, na níž článek začíná. Všechny tyto údaje jsou od sebe odděleny čárkami a mezerami, za číslem stránky se vyskytuje tečka. Záznam pokračuje názvem článku, na dalším řádku jsou pak uvedeny odkazy na příslušné citace ve Zlaté knize a rovněž na úplný text článku.

Pro názornost uveďme příklad záznamu časopiseckého článku:

PAC, 1972, 31, 577. Manual of Symbols and Terminology for Physicochemical Quantities and Units, Appendix II: Definitions, Terminology and Symbols in Colloid and Surface Chemistry
[218 entries] [Full text]

Citace záznamů na stránkách se zobrazením jednotlivých termínů a příslušných výkladů jsou uváděny slovně, není užito číselných odkazů (viz obr. 58). Slouží zároveň jako hypertextové odkazy na stránky se zobrazením podrobnějších

údajů o záznamu (včetně identifikátoru DOI, je-li dané publikaci přidělen), ale také odkazů na další termíny, resp. hesla, která citují týž dokument.

Source:

Green Book, p. 62

PAC, 1997, 69, 1007 (*Reference value standards and primary standards for pH measurements in D2O and aqueousorganic solvent mixtures: New accessions and assessments (Technical Report)*) on page 1007

PAC, 1996, 68, 957 (*Glossary of terms in quantities and units in Clinical Chemistry (IUPAC-IFCC Recommendations 1996)*) on page 986

See also:

PAC, 2002, 74, 2169 (*Measurement of pH. Definition, standards, and procedures (IUPAC Recommendations 2002)*) on page 2169

PAC, 1990, 62, 2167 (*Glossary of atmospheric chemistry terms (Recommendations 1990)*) on page 2205

PAC, 1984, 56, 567 (*Physicochemical quantities and units in clinical chemistry with special emphasis on activities and activity coefficients (Recommendations 1983)*) on page 569

Obr. 58: Příklad použitých a dalších zdrojů

7.4.2 Bibliografické údaje o knize

Podrobné informace o online verzi Zlaté knihy jsou k dispozici na samostatné stránce [291]; u jednotlivých termínů, resp. hesel jsou pak vždy uvedeny konkrétní údaje vztahující se ke knize jako k celku, ale také k heslům samotným (obr. 59). Zlatá kniha jako celek je doplněna identifikátorem DOI, rovněž jednotlivým termínům je přidělen tento identifikátor.

Cite as:

IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: <http://goldbook.iupac.org> (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook.

Last update: 2009-07-28; version: 2.1.3.

DOI of this term: doi:10.1351/goldbook.P04524.

Original PDF version: <http://www.iupac.org/goldbook/P04524.pdf>.

The PDF version is out of date and is provided for reference purposes only. For some entries, the PDF version may be unavailable.

Obr. 59: Vzor záznamu konkrétního termínu ve Zlaté knize

Ve zdrojovém kódu stránek, které tvoří součást webové prezentace Zlaté knihy, se setkáváme s využitím prvku „title“, v něm je uveden výraz „IUPAC Gold

Book“, po mezeře, pomlčce a další mezeře následuje výraz označující příslušné heslo, popř. jiné upřesnění obsahu stránky (obr. 60). Další prvky meta nejsou k zachycení bibliografických údajů využity.

```
<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org
/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"><head><meta
http-equiv="Content-Type" content="text/html; charset=UTF-8"
/><title>IUPAC Gold Book - pH</title><meta http-equiv="content-
language" content="en" /><link rel="stylesheet" type="text/css"
```

Obr. 60: Část zdrojového kódu stránky z webové prezentace Zlaté knihy

7.5 Zelená kniha – Množství, jednotky a symboly ve fyzikální chemii

Online verze Zelené knihy je k dispozici ve formátu PDF [35], informační stránka s údaji o knize je pak dostupná ve formátu HTML [292].

7.5.1 Záznamy použité literatury a citace

Záznamy použité literatury jsou uvedeny v samostatné kapitole (č. 9) nazvané *Záznamy (References)* – ta je umístěna v závěrečné části knihy. Kapitola tvoří celkem 75 položek, přičemž prvních sedm položek je součástí podkapitoly 9.1 Primární zdroje (*9.1 Primary sources*), dalších 40 položek tvoří obsah podkapitoly 9.2 Zdroje IUPAC (*9.2 IUPAC references*) a zbývajících 26 položek je uvedeno pod hlavičkou podkapitoly 9.3 Další zdroje (*9.3 Additional references*). S výjimkou primárních zdrojů, z nichž je v položce č. 1 zahrnuto jedenáct záznamů, v položce č. 2 čtyři záznamy a v položce č. 5 čtrnáct záznamů, jsou v jednotlivých položkách vždy zachyceny bibliografické údaje reprezentující jeden dokument.¹⁵⁷ Celkový počet záznamů uvedených v Zelené knize tedy dosahuje čísla 101.

Jak dokládá příklad záznamu časopiseckého článku, údaje o autorech jsou uváděny v invertované podobě, jména jsou zkracována na iniciály a v případě vyššího počtu autorů jsou údaje o jednotlivých autorech odděleny čárkami (s tím se ovšem nesetkáváme ve výše uvedeném příkladu) s tím, že jména posledních dvou autorů jsou oddělena spojkou „and“. Po čárce a mezeře následuje název článku, po další čárce a mezeře je kurzivou uveden zkrácený název časopisu, po mezeře pak následuje ročník časopisu zvýrazněný tučným písmem, poté je v uvozovkách

¹⁵⁷ Výjimku tvoří pouze záznam č. 4, v němž jsou uvedeny údaje o další dostupné podobě dokumentu; tyto údaje je však možné považovat za poznámku k prvnímu uvedenému záznamu.

prezentován rok vydání, který je od předchozích i následujících údajů rovněž oddělen mezerami. Jako poslední je uveden stránkový rozsah článku, tj. přesná lokace článku v daném ročníku časopisu. Záznam je ukončen tečkou.

Příklad záznamu časopiseckého článku:

Porter, H.Q. and Turner, D.W., A Descriptive Classification of the Electron Spectroscopies, *Pure Appl. Chem.* **59** (1987) 1343-1406.

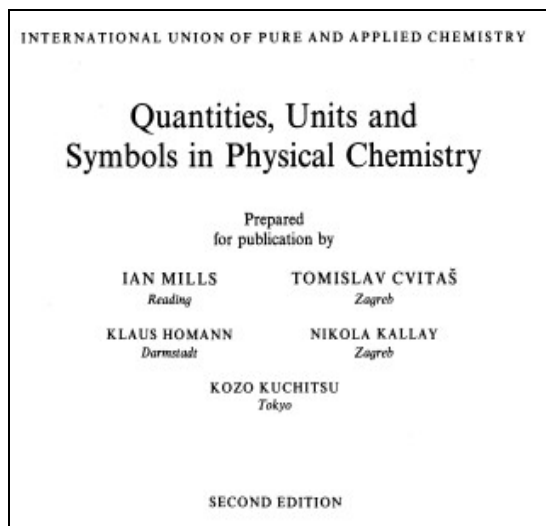
The *Manual of Symbols and Terminology for Physicochemical Quantities and Units* [1.a], to which this is a direct successor, was first prepared for publication on behalf of the Physical Chemistry Division of IUPAC by M.L. McGlashan in 1969, when he was chairman of the Commission on Physicochemical Symbols, Terminology and Units (I.1). He made a substantial contribution towards the objective which he described in the preface to that first edition as being 'to secure clarity and precision, and wider agreement in the use of symbols, by chemists in different countries, among physicists, chemists and engineers, and by editors of scientific journals'. The second edition of the manual prepared for publication by M.A. Paul in 1973 [1.b], and the third edition prepared by D.H. Whiffen in 1979 [1.c], were revisions to take account of various developments in the *Système International d'Unités (SI)*, and other developments in terminology.

Obr. 61: Příklad abecedně-číselných citací v textu

Pokud se týká citací, jsou v textu uváděny v hranatých závorkách. Jednotlivé záznamy jsou přitom reprezentovány buď jenom číselnými odkazy (není-li v seznamu zdrojů v položce uvedeno více záznamů zároveň), nebo číselně-abecedními odkazy (jedná-li se o některý ze tří případů, kdy je součástí jedné položky více záznamů, viz obr. 61).

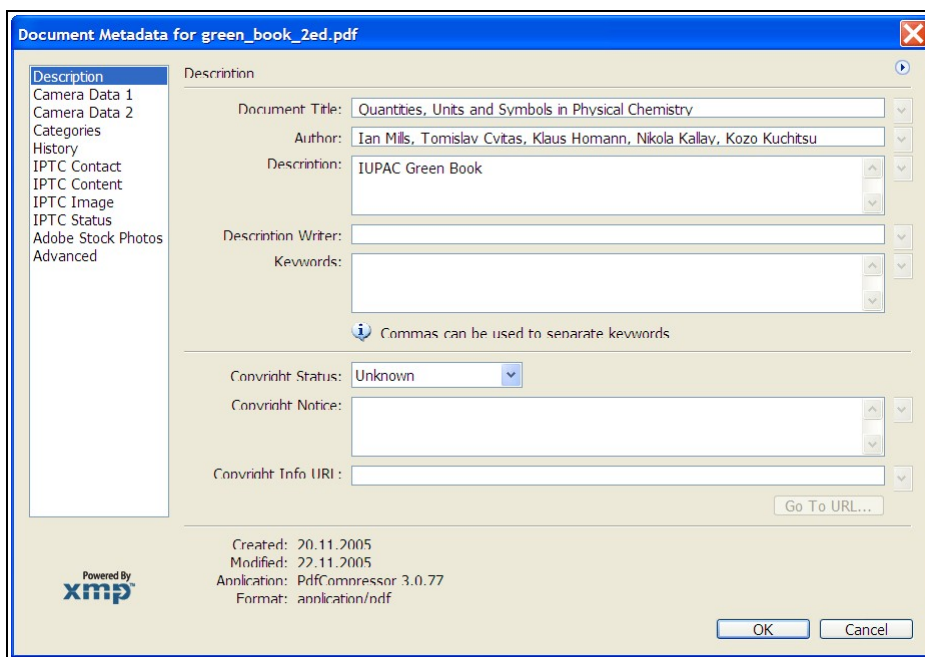
7.5.2 Bibliografické údaje o knize

Analyzovaná kniha je k dispozici ve formátu PDF, bibliografické údaje jsou uvedeny standardně na titulní straně a jejím rubu, resp. v elektronické podobě na titulní straně (viz obr. 62) a straně následující.



Obr. 62: Údaje na titulní straně Zelené knihy

Základní údaje o knize jsou uvedeny také ve vlastnostech dokumentu ve formátu PDF (obr. 63). Jedná se o název, o autory a dále o stručný název knihy – IUPAC Green Book. Ten je uveden v poli Popis (*Description*).



Obr. 63: Bibliografické údaje o Zelené knize ve vlastnostech dokumentu ve formátu PDF

7.6 Červená kniha – Nomenklatura anorganické chemie

Online verze Červené knihy je dostupná ve formátu PDF [36]. K dispozici je také webová stránka s informacemi o Červené knize [293].

7.6.1 Záznamy použité literatury a citace

Záznamy použité literatury jsou uváděny na konci každé kapitoly Červené knihy, celkem se s nimi setkáváme na jedenácti místech knihy. Záznamy jsou vždy předznamenány výrazem „References“, v řadě případů jsou však v rámci jednotlivých položek uváděny kromě záznamů použité literatury i poznámky, ve výjimečných případech (z celkového počtu 94 položek, viz tab. 1, se jedná o tři položky) jsou jako položky uváděny pouze poznámky (v jednom z výše zmíněných tří položek se jedná o poznámku doplněnou o odkaz na jiný záznam uvedený ve zdrojích).

V tab. 44 je uveden přehled počtu záznamů v jednotlivých kapitolách – první číslo vždy udává počet položek v části „References“, druhé číslo pak skutečný počet záznamů. Důvodem je jednak výše zmíněná skutečnost, že jsou v určitých případech jako jednotlivé položky uváděny i samostatné poznámky, jednak je potřeba

brát v úvahu, že v rámci některých položek je uvedeno více záznamů použité literatury (v takovém případě jsou jednotlivé záznamy odděleny středníkem a mezerou). Záznamy, u nichž je v závorce uvedeno stručné označení barevné knihy, popř. s upřesněním předpokládaného data vydání nové verze knihy, jsou pro naše účely pokládány vždy za jeden záznam, byť z pohledu dnešního čtenáře knihy se již může jednat o údaj o již vydaném novém dokumentu (viz obr. 64). Celkem se tedy v Červené knize setkáváme se 115 záznamy.

Tab. 44: Záznamy v Červené knize

Kapitola	Počet položek (počet všech záznamů)
IR-1	30 (39)
IR-2	6 (9)
IR-3	8 (8)
IR-4	5 (5)
IR-5	4 (4)
IR-6	3 (6)
IR-7	2 (3)
IR-8	4 (4)
IR-9	15 (17)
IR-10	6 (6)
IR-11	11 (14)
Celkem	94 (115)

3. *Compendium of Macromolecular Nomenclature*, ed. W.V. Metanomski, Blackwell Scientific Publications, Oxford, 1991. (The Purple Book. The second edition is planned for publication in 2005).

Obr. 64: Příklad záznamu s poznámkou o zkráceném označení dokumentu a plánovaném novém vydání

V žádném ze záznamů není uvedeno URL, nesetkáváme se ani s identifikátorem DOI či jiným náznakem odkazu na online zdroj. Řada citovaných zdrojů je však kromě standardní tištěné podoby již dostupná také prostřednictvím internetu, např. články publikované v časopise *Pure and Applied Chemistry*.

Příklad záznamu časopiseckého článku dokládá, že autorské údaje jsou uváděny v neinvertované podobě (začínají iniciálami), název časopisu je prezentován ve zkrácené podobě a je psán kurzivou, ročník je zvýrazněn tučně, číslo stránky, na níž článek začíná, je zachyceno běžným písmem. Tyto údaje jsou od sebe odděleny čárkami a mezerami. Posledním údajem záznamu je rok vydání (je uzavřen do kulatých závorek). Za tímto záznamem následoval v rámci jedné položky záznam

další, proto je v jeho závěru uveden středník (v případech, kdy se jedná o jediný záznam, je užívána tečka).

Příklad záznamu časopiseckého článku:

L.B. Guyton de Morveau, *J. Phys.*, **19**, 310 (1782);

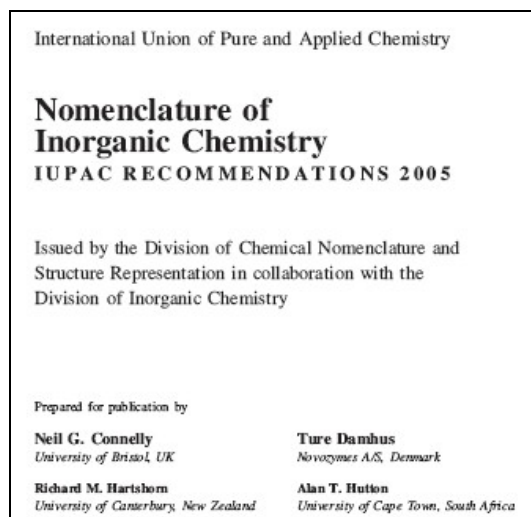
Na záznamy použité literatury je v textu odkazováno číselnými odkazy, které jsou uvedeny formou horního indexu (viz obr. 65).

composition. However, almost as soon as the true science of chemistry was established a 'system' of chemical nomenclature was developed by Guyton de Morveau in 1782.¹ Guyton's statement of the need for a 'constant method of denomination, which helps the intelligence and relieves the memory' clearly defines the basic aims of chemical nomenclature. His system was extended by a joint contribution² with Lavoisier, Berthollet, and de Fourcroy and was popularized by Lavoisier.³ Later, Berzelius championed

Obr. 65: Příklad citací použité literatury z kapitoly „IR-1 General Aims, Functions and Methods of Chemical Nomenclature“ z Červené knihy

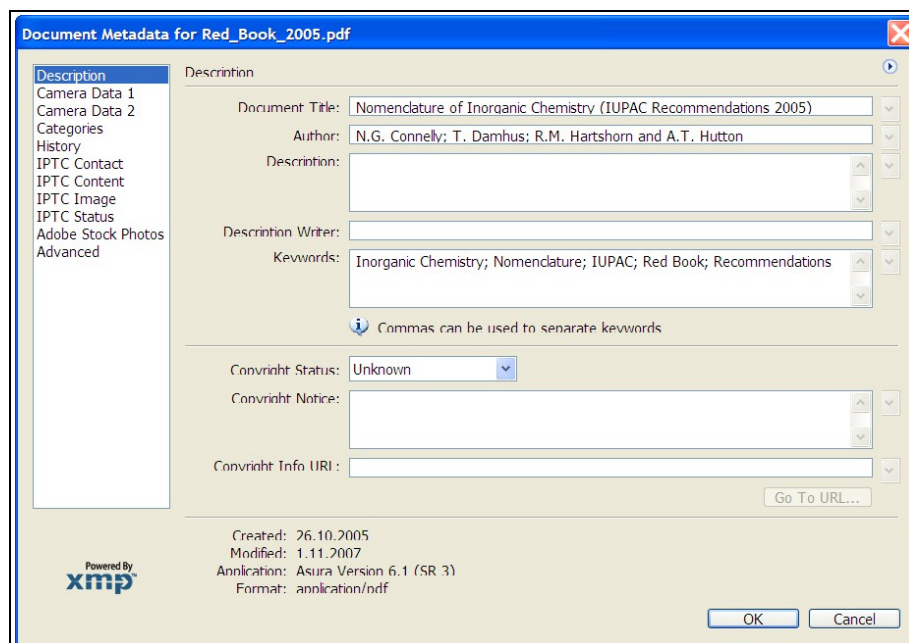
7.6.2 Bibliografické údaje o knize

Vzhledem k tomu, že se jedná o verzi publikace ve formátu PDF, jsou bibliografické údaje o knize uvedeny standardně na titulní straně a jejím rubu, resp. v elektronické podobě na titulní straně (obr. 66) a straně následující.



Obr. 66: Údaje na titulní straně Červené knihy

Ve vlastnostech dokumentu ve formátu PDF je taktéž možné zjistit základní údaje o dokumentu, konkrétně název a autory (obr. 67).



Obr. 67: Bibliografické údaje o Červené knize ve vlastnostech dokumentu ve formátu PDF

7.7 Modrá kniha – Nomenklatura organických sloučenin

Online verze Modré knihy je dostupná ve formátu HTML [37]. K dispozici jsou obě vydání Modré knihy – první z roku 1979 a druhé z roku 1993.

7.7.1 Záznamy použité literatury a citace

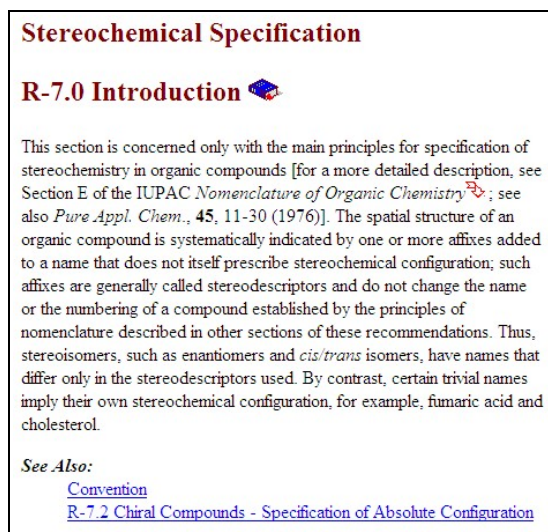
Online verze Modré knihy neobsahuje samostatnou část se záznamy použité literatury, je však možné si je (na samostatné HTML stránce) zobrazit v okamžiku, kdy se uživatel setká s citací v textu (není užívána metoda číselných odkazů, viz obr. 68) nebo je přímo v textu uveden záznam (viz obr. 69).

International Union of Pure and Applied Chemistry, Organic Chemistry Division, Commission on Nomenclature of Organic Chemistry, *Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F, and H*, 1979 Edition, J. Rigaudy and S. Klesney, eds, Pergamon Press, Oxford, 1979, 559 pp.

Obr. 68: Odkazovaný záznam

Výhodou v tomto případě je hypertextové propojení citací se záznamy, nevýhodou nemožnost zobrazení všech záznamů použité literatury najednou.

Z tohoto důvodu zde není uveden přesný počet záznamů vyskytujících se v online verzi Modré knihy.



Obr. 69: Příklad citace a záznamu v textu

7.7.2 Bibliografické údaje o knize

V zápatí domácí stránky Modré knihy jsou stručně uvedeny údaje o online verzi této knihy – je zpracována společností Advanced Chemistry Development, Inc.,¹⁵⁸ se svolením IUPAC. Jako zdrojové dokumenty jsou uvedeny dva tituly:

- *Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F, and H*, Pergamon Press, Oxford, 1979. Copyright 1979 IUPAC.
- *A Guide to IUPAC Nomenclature of Organic Compounds (Recommendations 1993)*, 1993, Blackwell Scientific publications, Copyright 1993 IUPAC.

Uživatelům, kteří potřebují Modrou knihu citovat, je doporučeno citovat tyto zdroje (tj. není doporučeno citovat online verzi Modré knihy).

Pokud se týká metadat, je v hlavičce domácí stránky online verze Modré knihy využit jednak prvek „title“ (jeho hodnotou je výraz „IUPAC Nomenclature“), jednak jsou využity prvky „meta“, konkrétně s atributy „description“ a „keywords“. Jako hodnota atributu „content“ v prvním zmíněném prvku jsou uvedeny bibliografické údaje týkající se původních tištěných předloh (obr. 70). Na dalších stránkách webové prezentace je v prvku „title“ uváděn název konkrétní kapitoly či podkapitoly, s prezentací bibliografických údajů v prvku „meta“ se již nesetkáváme.

¹⁵⁸ Jsou též uvedeny údaje o copyrightu – rok 1997 a název zpracovatelské společnosti.


```

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<html>
<HEAD>
<TITLE>IUPAC Nomenclature</TITLE>

<meta name="description" content="The HTML version of IUPAC
&quot;Blue Book&quot; Nomenclature of Organic Chemistry, Pergamon
Press, Oxford, 1979 and A Guide to IUPAC Nomenclature of Organic
Compounds (Recommendations 1993), 1993, Blackwell Scientific
publications.">

<meta name="keywords" content="IUPAC Rules, IUPAC Recommendations,
IUPAC Nomenclature, Chemical Nomenclature, Organic chemistry,
Organometallic compounds, Stereochemistry, Natural products, guide,
nomenclature, IUPAC, naming, Guide to IUPAC Nomenclature of Organic
Compounds, Blue Book, 1979, 1993">

</HEAD>

```

Obr. 70: Část zdrojového kódu stránky z webové prezentace Modré knihy

7.8 Fialová kniha – Makromolekulární nomenklatura

Fialová kniha je v online podobě dostupná ve dvou vydáních – v prvním [294] a ve druhém [38]. V následujících pasážích se zaměříme na vydání druhé, tj. novější.

V online verzi druhého vydání Fialové knihy lze identifikovat celkem 21 kapitol a dále přílohu, kterou tvoří bibliografie souvisejících doporučení týkajících se nomenklatury biopolymerů. Sedm z těchto kapitol a příloha nejsou k dispozici online, zbylých šestnáct kapitol je k dispozici ve formátu PDF (některé kapitoly jsou navíc k dispozici ve formátu HTML, kapitola 8 pak i ve formátu RTF). Jak vyplývá z vysvětlení uvedeného na domácí stránce online verze druhého vydání Fialové knihy, kapitoly této publikace tvoří oficiální doporučení vydaná v časopise Pure and Applied Chemistry. Přestože všechny články z tohoto časopisů tvořící součást Fialové knihy jsou k dispozici online, u některých kapitol nejsou ve druhém vydání Fialové knihy k dispozici URL odkazující na úplný text; tyto chybějící odkazy jsou uvedeny v tab. 45, zároveň je pro úplnost doplněn i identifikátor DOI.

Tab. 45: Doplněné odkazy na úplné texty kapitol Fialové knihy

Kapitoly	DOI	URL úplného textu
2.	doi:10.1351/pac19815303073 3	http://media.iupac.org/publications/pac/1981/pdf/5303x0733.pdf
3.	doi:10.1351/pac19896102021 1	http://media.iupac.org/publications/pac/1989/pdf/6102x0211.pdf
6.	doi:10.1351/pac19896104076 9	http://media.iupac.org/publications/pac/1989/pdf/6104x0769.pdf

Kapitoly	DOI	URL úplného textu
12.	konkrétní záznam článku není uveden, k dispozici je pouze odkaz na úvod k Fialové knize z roku 1987 a na nové dokumenty a aktualizace založené na nově publikovaných zprávách a doporučeních IUPAC	
14.	doi:10.1351/pac198557010149	http://media.iupac.org/publications/pac/1985/pdf/5701x0149.pdf
18.	doi:10.1351/pac198557101427	http://media.iupac.org/publications/pac/1985/pdf/5710x1427.pdf
21. (příloha)	online je k dispozici pouze příloha prvního vydání Fialové knihy, a to na adrese http://media.iupac.org/publications/books/pbook/PurpleBook-apd.pdf	

7.8.1 Záznamy použité literatury a citace

V tab. 46 jsou prezentovány počty položek, resp. počty všech záznamů uvedených v jednotlivých kapitolách Fialové knihy, jsou rovněž uvedeny příklady záznamů článků.

Tab. 46: Záznamy v kapitolách Fialové knihy

Kapitola	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
1.	0 (0)	samostatný záznam není k dispozici, jako příklad mohou být uvedeny bibliografické údaje o dokumentu přímo v textu: “Basic Definitions of Terms Relating to Polymers (1974)”, that appeared in <i>Pure Appl. Chem.</i> 1974 , <i>40</i> , 479–491.
2.	7 (8)	IUPAC Commission on Macromolecular Nomenclature. Basic Definitions of Terms Relating to Polymers 1974. <i>Pure Appl. Chem.</i> , 40 , 477 (1974).
3.	3 (4)	‘Basic Definitions of Terms Relating to Polymers 1974’, <i>Pure Appl. Chem.</i> 40 , 477-491 (1974).
4.	3 (4)	IUPAC. Glossary of Terms used in Physical Organic Chemistry. <i>Pure Appl. Chem.</i> 55 , 1281 (1983).
5.	5 (5)	International Union of Pure and Applied Chemistry. “Glossary of basic terms in polymer science”, <i>Pure Appl. Chem.</i> 68 , 2287–2311 (1996).
6.	11 (14)	‘Basic Definitions of Terms Relating to Polymers-1974’, <i>Pure Appl. Chem.</i> 40 , 477-491 (1974).
7.	39 (39)	D. Demus. <i>Liq. Cryst.</i> 5 , 75 (1989).
8.	15 (16)	J.M. Dealy, J. Rheol. 1984 28 181; 1995 39 253
9.	6 (6) – References 13 (13) – Bibliography	IUPAC. “Definitions of terms relating to crystalline polymers (IUPAC Recommendations 1988)” <i>Pure Appl. Chem.</i> 61 , 769–785 (1989).
10.	7 (8)	A. D. Jenkins, P. Kratochvíl, R. F. T. Stepto, U. W. Suter. “Glossary of basic terms in polymer science (IUPAC Recommendations 1996)”, <i>Pure Appl. Chem.</i> 68 , 2287–2311 (1996).
11.	0 (0)	záznamy zmíněny pouze v textu, např.: IUPAC Compendium of Chemical Terminology, 1987, p.199 ¹⁵⁹

Kapitola	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
12.	0 (0)	–
13.	11 (11) ¹⁶⁰	IUPAC. "Stereochemical definitions and notations relating to polymers 1980", <i>Pure Appl. Chem.</i> 53 , 733–752 (1981). Reprinted as Chapter 2 in IUPAC. <i>Compendium of Macromolecular Nomenclature</i> , Blackwell Scientific Publications, Oxford (1991).
14.	6 (8) ¹⁶¹	International Union of Pure and Applied Chemistry, "Stereochemical Definitions and Notations Relating to Polymers", <i>Pure Appl. Chem.</i> , 53 (3), 733-752 (1980).
15.	11 (12)	International Union of Pure and Applied Chemistry. "Glossary of Terms Used in Physical Organic Chemistry (Recommendations 1 9 8 2)". <i>Pure Appl. Chem.</i> 55 , 1281-1371 (1983).
16.	6 (6)	IUPAC. Commission on Macromolecular Nomenclature, "Source-based nomenclature for copolymers (rules approved 1985)", <i>Pure Appl. Chem.</i> 57 , 1427-1440 (1985). Reprinted as Chapter 7 in "Compendium".
17.	10 (10)	International Union of Pure and Applied Chemistry. "Source-Based Nomenclature for Copolymers" (1985). <i>Pure Appl. Chem.</i> 57 , 1427- 1440 (1985). Reprinted as chapter 7 in Ref. 1.
18.	8 (8)	Nomenclature of Regular Single—Strand Organic Polymers (1975), <i>Pure Appl. Chem.</i> 48 , 373 (1976).
19.	12 (12)	IUPAC. Definitions of terms relating to individual macromolecules, their assemblies, and dilute polymer solutions. <i>Pure Appl. Chem.</i> 61 , 211-241 (1989). Reprinted as Chapter 3 in ref. 1.
20.	7 (7)	"Source-based nomenclature for non-linear macromolecules and macromolecular assemblies", <i>Pure Appl. Chem.</i> 69 , 2511–2521 (1997).
21. (příloha, 1. vyd.)	10 (56) ¹⁶²	Abbreviations and symbols for the description of the conformation of polypeptide chains (tentative rules 1969). <i>Arch. Biochem. Biophys.</i> 145 , 405-421 (1971); ¹⁶³

Celkem je ve Fialové knize obsaženo 190 položek se záznamy, vlastních záznamů je pak 247 (připomeňme, že příloha v závěru knihy je na webu IUPAC

¹⁵⁹ V tomto případě se však nejedná však o záznam článku, ale o záznam monografie.

¹⁶⁰ Údaje o dostupnosti daného dokumentu také v jiných zdrojích (přetištění) jsou pro naše účely zjednodušeně považovány za poznámky tvořící integrální součást původního záznamu. Tento postup byl aplikován i v dalších obdobných případech.

¹⁶¹ Jako položka č. 4 je uveden záznam monografie „Nomenclature of Inorganic Chemistry“ (tj. Červené knihy); položka je pomocí písmen a až g dále členěna, jedná se však vždy o odkazy na konkrétní pravidla či tabulky z této monografie. Položka č. 4 je pro zjednodušení zachycena jako jeden záznam.

¹⁶² Upozornění na opravy nejsou pokládány za nový dokument, není-li jejich záznam uveden samostatně.

¹⁶³ Za záznamem následuje další záznam (v rámci jedné položky), proto tento příklad neuvádí tečku na konci záznamu – záznam ukončuje středník.

k dispozici pouze pro první vydání knihy, proto byl v tomto případě k analýze využit tento dostupný dokument).

Rozdíly v záznamech časopiseckých článků do určité míry odrážejí skutečnost, že se jedná o články z různých ročníků časopisu Pure and Applied Chemistry. Zatímco v novějších ročnících je např. uplatňováno uvádění ročníků tučným písmem, ve starších ročnících se s užitím tučného písma nesetkáváme; obdobně je tomu u názvů časopisů – ve starších ročnících byly zvýrazněny podtržením, v novějších je užitá kurziva. Z dalších rozdílných znaků lze jmenovat způsob uvádění údajů o autorské odpovědnosti (v některých případech je jako korporativní autor uváděn IUPAC, jinde je název organizace rozepsán, autorské údaje navíc nejsou vždy uvedeny), o názvu (setkáváme se s uzavřením do různých uvozovek, ale i s případy bez závorek) či o lokalizaci článku v daném ročníku časopisu (uvádění první strany článku či celého stránkového rozsahu).

Pokud se týká uvádění citací v textu, setkáváme se ve Fialové knize s různými variantami (viz tab. 47). Téměř vždy (kromě 1. a 11. kapitoly) jsou odkazy uváděny číselně a jsou uzavřeny do závorek, z toho v jedenácti případech do závorek kulatých a v sedmi případech do závorek hranatých. V osmi případech se setkáváme s uváděním číselných údajů bez slovního doprovodu (v případě shrnutí více citací zároveň je pouze doplněn spojovník), v devíti případech pak s uvedením vysvětlujícího výrazu (např. „ref.“ a „Ref.“, v plurálu „refs.“ a „Refs“.).

Tab. 47: Citace z jednotlivých kapitol Fialové knihy

Kapitola	Příklad citace
1.	–
2.	(1)
3.	(References 1-3)
4.	(ref. 1)
5.	[4]
6.	(Reference 3)
7.	(see refs. 1-39)
8.	(See, for example, ref. 11.)
9.	[3]
10.	[1]
11.	v textu jsou zmíněny pouze celé záznamy
12.	[3]
13.	[3]
14.	[Ref. 3a]
15.	(Ref. 2)
16.	(Ref. 4)

Kapitola	Příklad citace
17.	(Refs. 1 – 5)
18.	(Ref. 4)
19.	(2, 3)
20.	[1-3]
21. (příloha, 1. vyd.)	–

Pokud se týká oddělování záznamů v rámci položky, převažuje využití středníků a mezery, výjimečně se vyskytuje jiný způsob (ve 4. kapitole jsou jednotlivé záznamy uváděny na samostatných řádcích, a to za číselnou odrážkou, např. (i); ve 14. kapitole je využito oddělení středníkem a mezerou, přičemž další záznamy jsou předznamenány výrazy (a), (b) atd.; v 15. kapitole jsou odděleny tečkou a mezerou).

K propojení citací a záznamů nejsou ve většině případů použity hypertextové odkazy, výjimku tvoří např. HTML verze 5. kapitoly [295].

7.8.2 Bibliografické údaje o knize

Fialová kniha je složena z jednotlivých článků publikovaných v časopise *Pure and Applied Chemistry*, bibliografické údaje o nich jsou uvedeny jednak v rámci rozcestníku k úplným textům kapitol (obr. 71), jednak v jednotlivých úplných textech.

<u>Terminology Section</u>
1. Glossary of Basic Terms in Polymer Science (IUPAC Recommendations 1996). <i>Pure Appl. Chem.</i> 68 , 2287-2311 (1996) > www.iupac.org/reports/1996/6812jenkins/index.html
2. Stereochemical Definitions and Notations Relating to Polymers (IUPAC Recommendations 1980). <i>Pure Appl. Chem.</i> 53 , 733-752 (1981) > Purple Book I, Ch. 2.
3. Definitions of Terms Relating to Individual Macromolecules, their Assemblies, and Dilute Solutions (Recommendations 1988). <i>Pure Appl. Chem.</i> 61 , 211-241 (1989) > Purple Book I, Ch. 3.
4. Basic Classification and Definitions of Polymerisation Reactions (IUPAC Recommendations 1994). <i>Pure Appl. Chem.</i> 66 , 2483-2486 (1994) [full text - pdf 249 kB]
5. Definitions relating to stereochemically asymmetric polymerizations (IUPAC Recommendations 2001). <i>Pure Appl. Chem.</i> 74 , 915-922 (2002) > www.iupac.org/publications/pac/2002/7406/7406x0915.html

Obr. 71: Bibliografické údaje vztahující se k jednotlivým částem Fialové knihy

V prvcích „meta“ v hlavičce domácí stránky Fialové knihy (rozcestníku k jejím jednotlivým částem) jsou uvedena pouze klíčová slova, v prvku „title“ pak název organizace, nikoliv samotné knihy (obr. 72).

```
<HTML>
<HEAD>
  <META NAME="GENERATOR" CONTENT="Adobe PageMill 2.0 Mac">
  <TITLE>International Union of Pure and Applied Chemistry</TITLE>
  <META NAME="Keywords" CONTENT="IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC
IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC IUPAC
IUPAC IUPAC IUPAC IUPAC Chemistry Chemistry Chemistry Chemistry
Chemistry Chemistry Chemistry Chemistry Chemistry Chemistry
Chemistry Chemistry Chemistry Chemistry Chemistry Chemistry
Standards Standards Standards Standards Standards Standards
Standards Standards Standards Standards Standards Standards
Standards Standards Standards Standards">
</HEAD>
```

Obr. 72: Část zdrojového kódu domácí stránky webové prezentace Fialové knihy

7.9 Oranžová kniha – Analytická nomenklatura

Online verze třetího vydání Oranžové knihy [39] je prezentována částečně formou HTML stránek (v rámcích), ty však převážně slouží jako rozcestník k jednotlivým úplným textům ve formátu PDF.

7.9.1 Záznamy použité literatury a citace

Tab. 48 podává přehled o počtu položek, resp. počtu všech záznamů v jednotlivých kapitolách Oranžové knihy, jsou opět uvedeny i příklady záznamů článků.

Tab. 48: Záznamy v kapitolách Oranžové knihy

Kapitola	Počet položek (počet všech záznamů) ¹⁶⁴	Příklad záznamu (článek)
1.	4 (4)	J.Cesario, N.N.Greenwood, H.S.Peiser: Table of Standard Atomic Weights Abridged to Five Significant Figures Chem. International <u>15</u> 128-129 (1993)
2.	6 (6)	G.den Boef, A.Hulanicki: Recommendations for the usage of selective, selectivity and related terms in analytical chemistry PAC <u>55</u> (1983) 553-556
3.	15 (15)	Recommendation symbols for solution equilibria

¹⁶⁴ Záznamy nejsou číslovány (jsou od sebe odděleny prázdnými řádky), proto je v tomto případě počet položek shodný s počtem záznamů (z důvodu konzistence s ostatními daty uvedenými v tabulkách je uváděn i počet všech záznamů).

Kapitola	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
		PAC <u>18</u> (3) 437-464 (1969)
4.	1 (1)	Recommendations for terminology to be used with precision balances PAC <u>1</u> (1) 171-175 (1960)
5.	6 (6)	J.W.Stahl: Nomenclature of thermometric and enthalpimetric methods in chemical analysis. PAC <u>66</u> (12) 2487-2492 (1994)
6.	7 (7)	Recommended nomenclature for titrimetric analysis. PAC <u>18</u> (3) 427-436 (1969)
7.	6 (6)	Recommended nomenclature for automatic analysis. PAC <u>21</u> (1970) 527-531
8.	15 (15)	Recommendations for sign conventions and plotting of electrochemical data (Rules 1975) PAC <u>45</u> (2) 131-134 (1976)
9.	16 (16)	L. S. Ettre: Nomenclature for chromatography PAC <u>65</u> (4) 819-872 (1993)
10.	30 (30) ¹⁶⁵	A.M.Ure; L.R.P.Butler; R.O.Scott; R.Jenkins Preparation of Materials for Analytical Atomic Spectroscopy and Other Related Techniques PAC <u>60</u> 1461-1472 (1988)
11.	3 (3)	S.E.Braslavsky. K.N.Houk: Glossary of terms used in photochemistry PAC <u>60</u> (7) 1055-1106 (1988)
12.	3 (3)	J.F.Todd: Recommendations for nomenclature and symbolism for mass spectroscopy PAC <u>63</u> 1541-1566 (1991)
13.	1 (1)	není k dispozici, je uvedena pouze jedna poznámka se záznamem: "For a discussion of crystallographic nomenclature the International Tables for Crystallography Vol.A. (Edited by International Union of Crystallography, Kluwer Academic Publishers, Dordrecht, 1992) may be consulted."
14.	6 (6)	H.Kon: Recommendations for EPR/ESR nomenclature and conventions for presenting experimental data in publications. PAC <u>61</u> (12) 2195-2200 (1989)
15.7.2010	2 (2)	G.Svehla: Nomenclature of kinetic methods of analysis PAC <u>65</u> (10) 2291-2298 (1993)
16.	5 (5)	R.Van Grieken, M.De Bruin: Nomenclature for radioanalytical chemistry PAC <u>66</u> (122) 2513-2526 (1994)
17.	8 (8)	H.O.Porter, D.W.Turner: A descriptive classification of the electron Spectroscopies

¹⁶⁵ Zde se setkáváme s výjimkou, kdy je první část záznamů (konkrétně 21) číslována, zbytek záznamů již číslován není.

Kapitola	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
		PAC <u>59</u> (10) 1343-1406 (1987)
18.	19 (19)	W.Horwitz: Nomenclature of interlaboratory analytical studies PAC <u>66</u> (9) 1903-1911 (1994)
19.	14 (14)	R. Cornelius et al.: Sample collection guidelines for trace elements in blood and urine PAC <u>67</u> 1575-1608 (1995)

V Oranžové knize je obsaženo celkem 167 záznamů. V záznamech časopiseckých článků jsou typicky uváděny autorské údaje v neinvertované podobě (první, popř. i střední jméno je zkráceno na iniciály), v případě většího počtu autorů je užívána zkratka „et al.“ Po dvojtečce a mezeře následuje název článku, na dalším řádku pak zkrácený název časopisu (pro časopis Pure and Applied Chemistry není užíván zkrácený tvar Pure Appl. Chem., ale zkratka PAC), další údaje (ročník, stránkový rozsah a rok vydání) jsou odděleny pouze mezerami, ročník je zdůrazněn podtržením a rok vydání uzavřením do závorek. Lze konstatovat, že záznamy časopiseckých článků v Oranžové knize jsou konzistentní, u jednotlivých kapitol jsou uplatňována stejná pravidla pro jejich tvorbu.

Pokud se týká citací, nejsou uváděny v textu, ale naopak v závěrečných částech kapitol se záznamy je upřesněno, ke které části dané kapitoly se záznam zdroje vztahuje (viz obr. 73). Propojení s jednotlivými částmi kapitol není hypertextové, jedná se pouze o textové odkazy.

<p>1.9 References</p> <p><u>Source document for sections 1.2, 1.3, 1.6 and 1.7:</u> Quantities, Units and Symbols in Physical Chemistry (Green Book) Blackwell. Sci. Publ. Ltd. Oxford 1992</p> <p><u>Source document for section 1.4:</u> Bureau International des Poids et Mesures, Le Systeme, International d'Unités (SI), 6th French and English Edition, BI PM Sévres 1991</p>
--

Obr. 73: Ukázka ze seznamu zdrojů s uvedením vztahu k jednotlivým částem kapitoly

Ojedinele se další záznamy vyskytují i přímo v textu, např. v kapitole 11.1 (obr. 74). Ty do výše uvedeného počtu záznamů v Oranžové knize nejsou zahrnuty.

11. OTHER OPTICAL METHODS

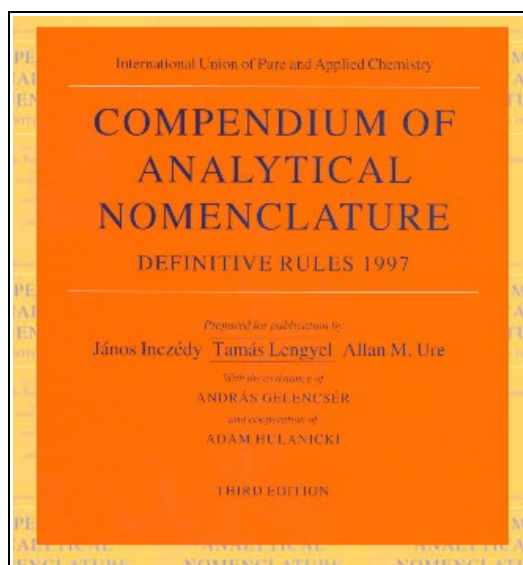
11.1 Introduction

This chapter contains terminology in the fields of photochemistry and light scattering and is based on 1) the glossary of terms used in the 'Glossary of Terms used in Photochemistry', Pure and Appl. Chem., 60 (1988) 1055, prepared by the Photochemistry Commission of the Organic Chemistry Division of IUPAC and 2) on Appendix II - Definitions, terminology and symbols in colloid and surface chemistry, Manual of Symbols and Terminology for Physicochemical Quantities and Units, PART 1.14: Light Scattering, published in Pure and Appl. Chem., 55 (1983) 931.

Obr. 74: Příklad záznamů v textu Oranžové knihy

7.9.2 Bibliografické údaje o knize

Bibliografické údaje o Oranžové knize jsou uvedeny formou obrázku obálky tištěného vydání (obr. 75), v levé dolní části stránky (v levém rámci) jsou doplněny údaje o původci webového sídla a o datu poslední aktualizace.



Obr. 75: Bibliografické údaje o tištěné verzi Oranžové knihy na obálce

Ve zdrojovém kódu hlavního rámce domácí stránky online verze Oranžové knihy je v prvku „title“ uvedeno „IUPAC Orange Book Cover“ (obr. 76), u dalších stránek (hlavních rámců) se vyskytují názvy souborů ve formátu HTML, např. „TOC_cha10.html“, popř. výraz „Untitled“.

```

<!doctype html public "-//w3c//dtd html 4.0 transitional//en">
<html>
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=iso-
8859-1">
  <meta name="GENERATOR" content="Mozilla/4.75C-CCK-MCD {C-UDF;
EBM-APPLE} (Macintosh; U; PPC) [Netscape]">
  <title>IUPAC Orange Book Cover</title>
</head>
<body background="OBcoverfade.jpg">

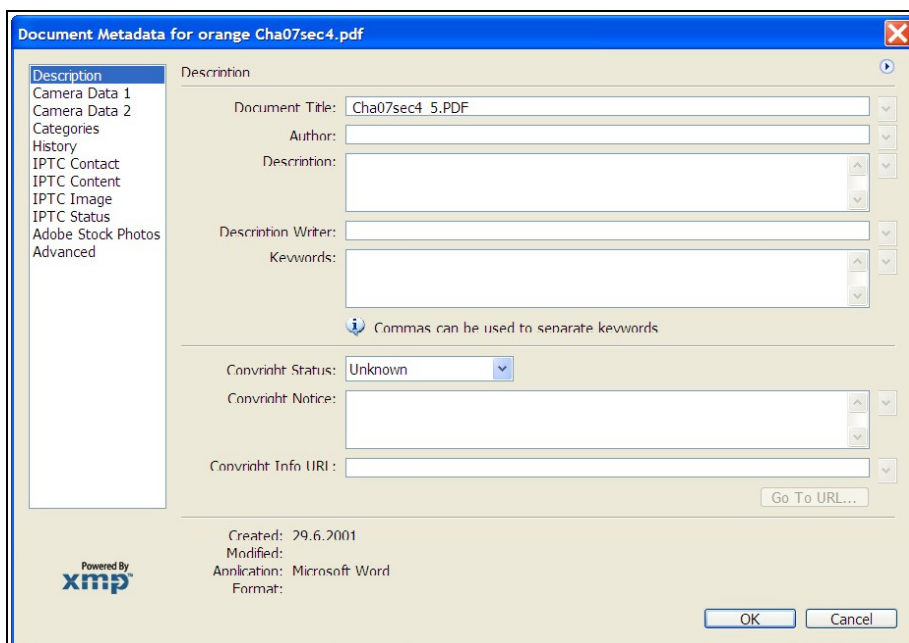
<center><img SRC="OrangeBookCover.jpg" BORDER=0 NATURALSIZEFLAG="2"
height=329 width=300 align=BOTTOM></center>

</body>
</html>

```

Obr. 76: Zdrojový kód hlavního rámce domácí stránky online verze Oranžové knihy

V jednotlivých souborech s kapitoly knihy ve formátu PDF jsou ve vlastnostech dokumentu bibliografické údaje uváděny různým způsobem, v řadě případů nejsou uváděny vůbec, resp. jako název dokumentu je uveden název souboru (obr. 77).



Obr. 77: Metadata v souboru ve formátu PDF s kapitolou 7.4 z Oranžové knihy

7.10 Stříbrná kniha – Nomenklatura a symboly v klinické chemii

V rámci webové prezentace organizace IUPAC jsou k dispozici pouze stručné informace o knize, nikoliv však samotný obsah [296]. Tato kniha proto v této práci není předmětem podrobnější analýzy.

7.11 Bílá kniha – Biochemická nomenklatura a související dokumenty

V rámci online prezentace Bílé knihy [40] jsou k dispozici pouze některé části této knihy, a to ve formátu PDF (reprezentující původní tištěnou podobu) nebo ve formátu HTML (v tomto případě se často jedná o revidované verze původních tištěných dokumentů).

7.11.1 Záznamy použité literatury a citace

Tab. 49 poskytuje názorný přehled o počtu položek, resp. všech záznamů, nechybějí příklady záznamů časopiseckých článků.

Tab. 49: Záznamy v kapitolách Bílé knihy

Kapitola, resp. stránkový rozsah	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
1-18	0 (0) ¹⁶⁶	J. M. Bijvoet, A. F. Peerdeman and A. J. van Bommel, <i>Nature</i> 168 , 271 (1951);
19-26	5 (8) ¹⁶⁷	International Union of Biochemistry, International Union of Pure and Applied Chemistry, Commission on Biochemical Nomenclature, "Nomenclature of Carotenoids", <i>Pure Appl. Chem.</i> , 41 , 405-431 (1975).
27-30	1 (2) ¹⁶⁸	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), and Nomenclature Commission of IUB (NC-IUB), Newsletter 1980, <i>Eur. J. Biochem.</i> , 1980, 104 , 321-322
31-36	5 (5)	Alberly, R. A. (1992) <i>Biophys. Chem.</i> 43 , 239-254.
37-38	0 (0) ¹⁶⁹	–
39-69	37 (129)	International Union of Pure and Applied Chemistry (IUPAC), Definitive Rules for the Nomenclature of Amino Acids (1960) <i>J. Amer. Chem. Soc.</i> 82 , 5575-5577.
70-72	7 (40)	<i>J. Biol. Chem.</i> , 241 , 2491 (1966)
73-81	8 (11)	Cahn, R.S. (1964), <i>J. Chem. Educ.</i> 41 , 116.

¹⁶⁶ Záznamy jsou uváděny v poznámkách pod čarou v textu, nikoliv v závěru kapitoly. Jsou ukončeny tečkou, v příkladu se vyskytuje středník z důvodu uvedení dalšího záznamu v téže poznámce.

¹⁶⁷ Vzhledem ke způsobu svého uvedení jsou v tomto případě záznamy uvedené po frázi typu „also in“ započítány jako samostatné úplné záznamy.

¹⁶⁸ Další záznamy se vyskytují v textu.

¹⁶⁹ Uvedená kapitola není k dispozici online, proto jsou ve sloupci uvedeny nulové hodnoty.

Kapitola, resp. stránkový rozsah	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
82-83	0 (0) ¹⁷⁰	<i>Biochem. J.</i> , 1975, 151 , 1-4
84-89	11 (23)	Schleifer, K. H. & Kandler, O. (1972) <i>Bacteriol. Rev.</i> 36 , 407-477.
90-92	2 (2)	B. Safer, <i>Eur. J. Biochem.</i> , 1989, 186 , 1-3.
93-95	4 (18)	Webb, E. C. (1964) <i>Lancet</i> 1 , 1110
96-106	6 (8) ¹⁷¹	International Union of Pure and Applied Chemistry (1981) Symbolism and terminology in chemical kinetics, 1980, <i>Pure Appl. Chem.</i> 53 , 753 - 771.
107-108	1 (1)	von Döhren, H. (1980) <i>Trends Biochem. Sci.</i> 5 (3), VIII.
109-114	17 (20)	Holley, R. W., <i>Progr. Nucl. Acid. Res. Mol. Biol.</i> 8 (1968) 37.
115-121	18 (33)	IUPAC Commission on Macromolecular Nomenclature (CMN), Stereochemical definitions and notations relating to polymers, <i>Pure Appl. Chem.</i> 53 , 733-752 (1981).
122-126	0 (0) ¹⁷²	–
127-148	27 (68)	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Polysaccharide nomenclature (Recommendations 1980), <i>Eur. J. Biochem.</i> , 126 , 439-441 (1982)
149-155	13 (44)	Hanson, K. R. (1966) <i>J. Am. Chem. Soc.</i> 88 , 2731.
156-157	9 (26)	Parthasarathy, R. & Eisenberg, E., Jr (1986) <i>Biochem. J.</i> 235 , 313-322
158-161	0 (0)	–
162-164	4 (7)	IUPAC Commission on Nomenclature of Organic Chemistry, Rules for the nomenclature of organic chemistry, section E, stereochemistry (recommendations 1974) <i>Pure Appl. Chem.</i> 45 , 11–30 (1976).
165-168	4 (7)	IUPAC Commission on Nomenclature of Organic Chemistry, Rules for the nomenclature of organic chemistry, section E, stereochemistry (recommendations 1974) <i>Pure Appl. Chem.</i> 45 , 11–30 (1976).
169-173	7 (33)	Mills, J. A. (1955) <i>Adv. Carbohydr. Chem.</i> 10 , 1-53
174-176	4 (12)	<i>Nomenclature of Regular Single-Strand Organic Polymers</i> (1976) <i>Pure Appl. Chem.</i> 48 , 375–385
177-179	8 (27)	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) Conformational nomenclature for five and six-membered ring forms of monosaccharides and their derivatives, Recommendations 1980. <i>Arch. Biochem. Biophys.</i> 207 , 469-472 (1981)
180-190	13 (39)	Mills, J. A. & Klyne, W. (1954) <i>Progr. Stereochem.</i> 1 , 181.
192-221	6 (17)	IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN). The Nomenclature of Steroids, Revised tentative rules, 1967. <i>Arch. Biochem. Biophys.</i> 136 , 13-35 (1970), amended 147 , 4-7 (1971)
222-225	0 (0) ¹⁷³	<i>Arch. Biochem. Biophys.</i> , 1974, 165 , 1-5
226-238	5 (5) ¹⁷⁴	<i>Pure Appl. Chem.</i> 11 , Nos. 1 and 2 (1965)

¹⁷⁰ V tomto případě se vyskytují pouze záznamy v textu (z nich je vybrán příklad), nevyskytují se však v závěru kapitoly, proto není uveden jejich počet.

¹⁷¹ Tato kapitola byla dále rozčleněna na podkapitoly, z nichž u některých byly uvedeny záznamy použité literatury. Protože se záznamy opakují, je uveden počet odkazů týkající se závěrečné části „References“.

¹⁷² Kapitola není k dispozici (odkaz na ni je nefunkční).

¹⁷³ Uvedeny jsou pouze záznamy v textu, z nich je vybrán příklad.

Kapitola, resp. stránkový rozsah	Počet položek (počet všech záznamů)	Příklad záznamu (článek)
239-241	17 (33)	Ames, S. R. (1979) <i>J. Nutr.</i> 109 , 2198-2204.
242-246	10 (33)	IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN). The Nomenclature of Steroids, Revised tentative rules, 1967, <i>Arch. Biochem. Biophys.</i> 136 , 13-35 (1970) amended 147 , 4-7 (1971)
247-251	18 (35)	IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC), Rules for the nomenclature of organic chemistry, Section E: Stereochemistry, Recommendations 1974, <i>Pure Appl. Chem.</i> 45 , 11-30 (1976)
252-255	16 (22)	Popják, G. & Cornforth, J. W. (1960) <i>Adv. Enzymol.</i> 22 , 281-335.
256-264	12 (18)	Nomenclature of Carbohydrates (1969), <i>Biochemistry</i> 10 , 3983-4004 (1971)
266-268	3 (14)	IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Nomenclature and symbols for folic acid and related compounds. <i>Arch. Biochem. Biophys.</i> 118 , 511-513 (1967)
269-271	9 (9)	International Union of Pure and Applied Chemistry (1960) <i>J. Am. Chem. Soc.</i> 82 , 5581.
272-277	7 (20)	IUPAC-IUB (1966), <i>Biochem. J.</i> 102 , 19
278-329	13 (13)	P. O'Carra and S. D. Killilea, <i>Tetrahedron Letters</i> , 4211 (1970).
331	1 (4) 2 (7) ¹⁷⁵	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), and Nomenclature Commission of IUB (NC-IUB), Newsletter 1981, <i>Arch. Biochem. Biophys.</i> , 1981, 206 , 458-462
332	1 (4) 1 (4) ¹⁷⁶	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), and Nomenclature Commission of IUB (NC-IUB), Newsletter 1981, <i>Arch. Biochem. Biophys.</i> , 1981, 206 , 458-462
333	1 (8) 1 (8) 3 (20)	IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), and Nomenclature Commission of IUB (NC-IUB), Newsletter 1984, <i>Arch. Biochem. Biophys.</i> , 1984, 229 , 237-245
334	4 (11) 2 (6) 3 (7) 3 (7)	Bergmann, M., Zervas, L. & Silberkweit, E. (1931) <i>Naturwissenschaften</i> 19 , 20.
335	2 (7) 1 (6)	IUPAC-IUBMB Joint Commission on Biochemical Nomenclature (JCBN), and Nomenclature Commission of IUBMB (NC-IUBMB), Newsletter 1992, <i>Arch. Biochem. Biophys.</i> , 1992, 294 , 322-325
336-340	0 (0) ¹⁷⁷	–
341-343	0 (0)	–
244-347	0 (0)	–

¹⁷⁴ V tomto případě se jedná spíše o poznámky doplněné o záznamy.

¹⁷⁵ Kapitola je složena ze dvou částí, je vybrán příklad z první části; s obdobnou situací se setkáváme v dalších kapitolách na s. 332-335, proto je v příslušných buňkách tabulky ve sloupci s počtem záznamů uvedeno více hodnot.

¹⁷⁶ Další záznam je uveden v textu.

¹⁷⁷ Tato kapitola a následující dvě kapitoly nejsou k dispozici online.

Celkem Bílá kniha obsahuje 351 položek, počet záznamů dosahuje čísla 881.¹⁷⁸

Záznamy z jednotlivých kapitol Bílé knihy se vzájemně značným způsobem liší, zjednodušeně lze konstatovat, že se shodují pouze v uvádění zkrácených názvů časopisů, a to kurzivou. Autorské údaje jsou zachyceny v invertované i neinvertované podobě, v případě více autorů jsou údaje o posledních dvou z nich spojeny výrazem „and“ nebo znakem „&“, názvy článků jsou uváděny pouze v některých případech, ročník je převážně zdůrazněn tučným písmem (není to však pravidlo), častěji se setkáváme s uváděním celého stránkového rozsahu než pouze první strany článku, roky vydání jsou často uváděny v závorkách, převážně na konci záznamu, opět se však nejedná o pravidlo, které by bylo dodržováno v celé Bílé knize.

V této knize také často nastává případ, kdy je vstupní prvek (název doporučení – článku apod.) uveden jenom na začátku dané položky, u dalších je pak již vynechán.

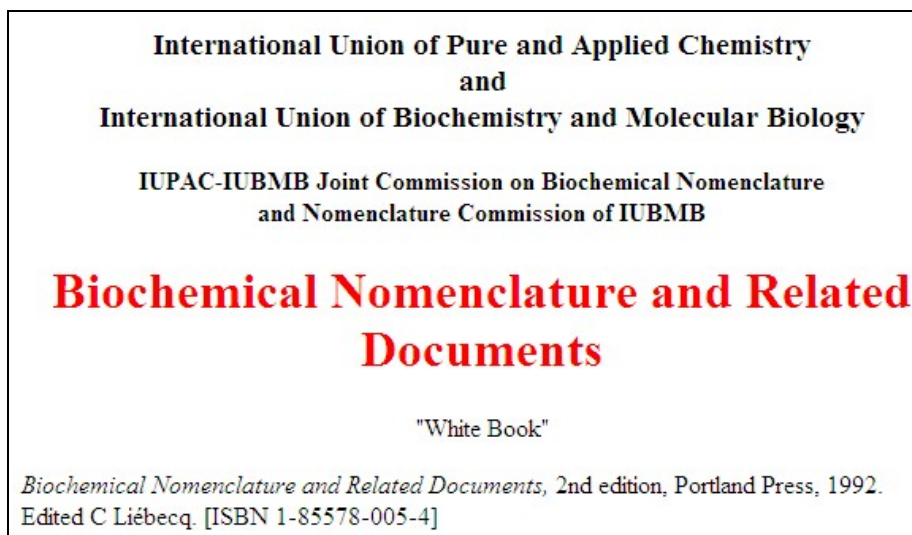
Citace jsou v textu kapitol uváděny různým způsobem – v případě, že se jedná o kapitolu dostupnou nejenom ve formátu PDF, ale také ve formátu HTML, jsou citace v textu provázány hypertextovým odkazem na příslušný záznam v seznamu zdrojů. Převažuje uvádění citací formou číselných odkazů, přičemž čísla jsou uzavřena do kulatých nebo hranatých závorek (v citaci je uváděno jedno číslo nebo rozsah čísel, v těchto případech je doplněn znak spojovníku nebo pomlčky). V některých případech je dané číslo předznamenáno výrazem „ref“. Ojedinele je uplatněn harvardský způsob citování prostřednictvím vstupního prvku záznamu a data vydání (např. v kapitolách na s. 73-81 a 90-92); vyskytuje se též uvádění záznamů zdrojů v poznámkách pod čarou (v kapitole na s. 1-18).

Pokud se týká uvádění většího počtu záznamů v rámci jedné položky, převažuje jejich oddělování středníkem a mezerou, zřídka je využita čárka a mezera (např. v kapitole na s. 162-164).

¹⁷⁸ Dodatky vyskytující se u některých kapitol (ale netvořící jejich integrální součást), další záznamy uvedené na domácí stránce online verze Bílé knihy ani stručná upozornění na dokumenty v hranatých závorkách nebyly započítány. Obdobně bylo naloženo i s výrazy typu „and elsewhere“, které neodkazují na konkrétní zdroj.

7.11.2 Bibliografické údaje o knize

Bibliografické údaje týkající se druhého tištěného vydání Bílé knihy jsou uvedeny v záhlaví HTML stránky, která plní funkci rozcestníku k jednotlivým částem této knihy (obr. 78).



Obr. 78: Bibliografické údaje týkající se tištěného vydání Bílé knihy

V prvcích „meta“ v hlavičce HTML stránky, která slouží jako rozcestník, jsou uvedeny pouze klíčová slova, popis a typ objektu, jako hodnota prvku „title“ je uveden zkrácený oficiální název Bílé knihy – „Biochemical Nomenclature Book“ (obr. 79) Další bibliografické údaje v hlavičce prezentovány nejsou. Analogicky je tomu u jednotlivých dalších stránek, v prvku „title“ je však typicky uváděn zkrácený název kapitol či podkapitol (obr. 80).

```
<html><head><title>Biochemical Nomenclature Book</title>
<base href="http://www.chem.qmul.ac.uk/iupac/bibliog/white.html">
<meta name="keywords" content="IUPAC IUBMB biochemical biochemistry
chemistry nomenclature bibliography translation">
<meta name="description" content="Bibliography of IUPAC and IUBMB
biochemical nomenclature recommendations, and translations">
<meta name="object type" content="document">
</head><body>
```

Obr. 79: Část zdrojového kódu domácí stránky webové prezentace Bílé knihy

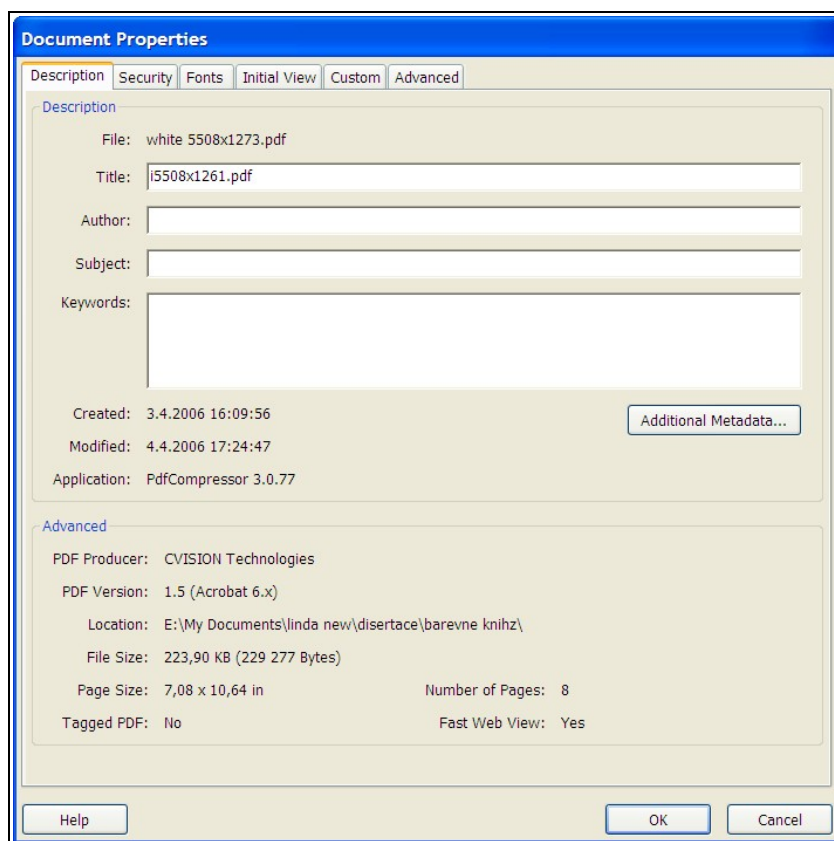
```

<html><head><title>Polynucleotide Conformation 1</title>
<base href="http://www.chem.qmul.ac.uk/iupac/misc/pnucl.html">
<meta name="keywords" content="IUPAC IUBMB biochemistry symbols
conformation polynucleotide nomenclature terminology">
<meta name="description" content="The full text of the IUPAC and
IUBMB abbreviations and symbols for the description of conformations
of polynucleotide chains">
<meta name="object type" content="document">
</head><body>

```

Obr. 80: Část zdrojového kódu jedné ze stránek webové prezentace Bílé knihy

Pokud se týká kapitol dostupných ve formátu PDF, bibliografické údaje jsou uváděny v záhlaví první stránky (případné poznámky týkající se původního vydání dokumentu v jiném časopise jsou zachyceny naopak v patičce). Ve vlastnostech dokumentu ve formátu PDF tyto údaje uváděny nejsou. V některých případech se setkáváme s uvedením názvu souboru jako názvu dokumentu (obr. 81).



Obr. 81: Bibliografické údaje o Bílé knize ve vlastnostech dokumentu ve formátu PDF

7.12 Dílčí závěry

Jako nejpodstatnější závěr lze uvést, že jednotlivé barevné knihy jsou zpracovány různým způsobem, se značnými odlišnostmi se lze setkat i v oblasti citování, tj. uvádění citací v textu a záznamů v seznamu zdrojů. Různorodost citačních stylů (v řadě případů spíše způsobů citování) je zřejmá již při porovnání záznamů článků (tj. jednoho konkrétního typu dokumentu). Jednou z příčin této různorodosti je samozřejmě různá doba vzniku jednotlivých barevných knih, další je skutečnost, že na rozdíl od časopisů vydávaných organizací IUPAC se podoba citací a záznamů neřídí konkrétními citačními instrukcemi či pokyny.

Vyšší stupeň standardizace v této oblasti lze pokládat za žádoucí, a to i s ohledem na plánované postupné zpracovávání barevných knih do strojově srozumitelné podoby (s využitím technologií založených na jazyce XML). Barevné knihy hrají v oblasti chemické terminologie a nomenklatury standardizační (sic!) úlohu, proto je pochopitelné, že na sebe vzájemně často odkazují (obecně lze říci, že totéž platí o všech publikacích organizace IUPAC). Při přípravě strojově srozumitelných verzí knih by proto za předpokladu sjednocení používaných citačních stylů, resp. citačního formátu, který by byl použit pro vygenerování citačního výstupu pro lidského uživatele, bylo možné jednotlivé barevné knihy vzájemně lépe automaticky provázat. Vzhledem k tomu, že celkový počet záznamů použité literatury v barevných knihách přesahuje číslo 1 600 a že dochází k aktualizaci obsahu jednotlivých knih, je výhodnost automatického provázání zcela zřejmá. Navíc budou-li k dispozici citační instrukce a budou-li dodržovány, bude díky zpracování barevných knih ve strojově srozumitelné podobě možné mj. automaticky získávat údaje pro vizualizaci vzájemných citačních vztahů.

Dalším význačným rysem je pak úzké spojení knih s časopisem *Pure and Applied Chemistry*, který je rovněž publikačním výstupem organizace IUPAC – vysoká míra závislosti barevných knih na tomto časopise (resp. článcích v něm publikovaných) je výrazná zejména u Fialové knihy, která představuje soubor článků vydaných právě v tomto časopise.

Citační praxe současných online verzí barevných knih navíc ukazuje potřebu sjednotit způsob zahrnutí odkazů na online zdroje (a jejich dílčí části), resp. jejich

identifikátory. V případě všech článků z časopisu Pure and Applied Chemistry je již možné využít odkazování prostřednictvím identifikátoru DOI, stejně tomu je i u online verze Zlaté knihy (i všech jejích dílčích součástí).

Pokud se týká uvádění bibliografických údajů o knihách samotných, lze konstatovat, že tyto údaje jsou převážně k dispozici v podobě využitelné pouze člověkem, v případě verzí ve formátu HTML nejsou uváděny podrobnější údaje formou metadat v hlavičce stránky, v případě verzí ve formátu PDF je tomu obdobně, pokud se týká těchto údajů ve vlastnostech dokumentu.

Je také potřeba poznamenat, že žádná z online verzí barevných knih v současné době uživatelům nenabízí možnost stažení bibliografických údajů v citačních formátech (např. RIS či BibTeX) nebo přímého importu těchto údajů do citačních manažerů (např. do citačního manažeru Zotero). Alespoň částečnou automatizaci při ukládání dat však lze aplikovat i v případě barevných knih, např. využít citačních manažerů, které primárně fungují jako nástroje na ukládání záložek a dokáží z webových stránek automaticky získat alespoň některé údaje, např. URL a název (patří mezi ně mj. Connotea).

Výsledky předložené analýzy budou v praxi využity výborem CPEP, a to konkrétně v následujícím období při přípravě strojově srozumitelné verze Červené a Zelené knihy, jejichž zpracování navazuje na již nyní dostupnou Zlatou knihu.

8. Zahrnutí citačních dat do XML formátů

8.1 Úvod

Tato kapitola navazuje zejména na zjištění z kapitol 3 a 4. Lze konstatovat, že vzrůstá potřeba formátu v jazyce XML, ve kterém by bylo možné zachytit citační data způsobem, který by umožňoval jejich následné efektivní využití pro účely různých informačních systémů.

XML formát pro zachycení citačních dat by měl nabídnout možnost efektivně (a do určité míry i kreativně) pracovat s daty v něm zachycenými. Na rozdíl od tradičně užívaných formátů typu RIS, EndNote či BibTeX by se nemělo jednat o formát primárně určený jako vstupní formát dat do konkrétního nástroje či systému. Hlavní výhodou XML formátu je jeho využitelnost pro uložení dat v libovolném informačním systému, který je založen na využití jazyka XML, případně je schopen importovat určitou část dat do svých datových struktur a zároveň celý záznam uložit v nezměněné XML podobě. Všechny další potřebné výstupy je pak možné zajišťovat podle konkrétních požadavků, součástí mohou být i výstupy do tradičně užívaných formátů či jiných XML formátů, s nimiž se setkáváme v oblasti digitálních knihoven a metadat. Předpokladem pro tvorbu těchto výstupů je vytvoření mapovací tabulky, v níž by byly prvky v navrženém formátu namapovány na odpovídající prvky v požadovaném výstupním formátu. V tomto okamžiku se již lze setkat s určitými ztrátami dat či nepřesnostmi, které vyplývají z architektury výstupního formátu, resp. z jeho omezení.

Data v XML formátu navíc mohou v případě potřeby být využita pro vytváření alternativních rozhraní digitálních knihoven. Součástí formátu mohou být také data (resp. metadata) týkající se vlastních citačních dat (mohou např. zachycovat způsob získání citačních dat, datum a čas jejich získání či údaje o jejich zdroji) – uvádění dat tohoto typu je zatím typické spíše pro metadatové formáty primárně užívané pro jiné než citační účely.

8.2 Doporučené postupy

Na základě dílčích závěrů z předchozích kapitol je možné formulovat doporučené postupy pro zahrnutí citačních dat do XML formátů z oblasti STM. Doporučené postupy jsou naznačeny formou jednotlivých zásad.

1. Využívat implicitní kódování XML, tj. UTF-8.

Využití UTF-8, tj. implicitního kódování jazyka XML¹⁷⁹, je vhodné zejména s ohledem na skutečnost, že pro oblast citování je charakteristická rozmanitost používaných znaků. Na rozdíl od dřívějších dob nejsou při zpracování počítačových dat používány například pouze znaky tvořící součást kódování ASCII. Je potřeba mít možnost vyjádřit řadu jiných znaků (např. znaků s diakritikou v české abecedě nebo znaků řecké, arabské či hebrejské abecedy). Právě pro tento účel lze kódování UTF-8 účelně využít. Navíc v sobě automaticky zahrnuje zmíněné kódování ASCII, je zpětně kompatibilní – prvních 128 znaků kódování UTF-8 jsou právě znaky ASCII.

2. Pracovat s relativně nízkým počtem prvků i atributů.

Díky nízkému počtu prvků a atributů lze přispět ke snížení pravděpodobnosti chyb vzniklých nepřesným či nekonzistentním způsobem jejich užití. Je-li totiž počet prvků dostatečně nízký, je možné si jednotlivé prvky v případě potřeby snadno zapamatovat a efektivně a konzistentně s nimi v praxi pracovat. Nízký počet prvků a atributů také přispívá ke snížení míry variability způsobu zachycení dat.

3. Co nejvíce strukturovat citační data, aby bylo následně možné efektivně provádět jejich potřebné transformace.

Úspěšnost využití XML formátu pro přenos dat závisí samozřejmě na přesnosti dat, resp. na hodnotách konkrétních prvků, popř. v XML formátu také atributů. Zvýšení přesnosti dat je možné docílit zavedením co nejvyšší granularity dat (s případnými doplňkovými pravidly pro tvorbu obsahu, tj. vyplňování hodnot prvků). V tomto směru jazyk XML neklade žádné omezení. Navíc nabízí nástroje,

¹⁷⁹ Spolu s kódováním UTF-8 je implicitním kódováním XML dokumentů také kódování UTF-16. Pokud kódování není uvedeno explicitně, procesory od sebe obě kódování dokáží automaticky rozeznat.

jak provádět kontrolu dat formalizovaně, konkrétně pomocí schémat. Těmito schémata mohou být DTD, XML Schema či Relax NG.

Formát pro zachycení citačních dat konstruován tak, aby byly jako hodnoty jednotlivých prvků uváděny pokud možno dále nedělitelné hodnoty údajů. Například máme-li ve formátu zachytit jméno autora složené z iniciály prvního jména, iniciály středního jména a celého příjmení, nebudou tyto hodnoty uvedeny v rámci jednoho prvku, jako je tomu např. ve formátech RIS či BibTeX (v něm je v rámci jednoho prvku uváděno i více autorů zároveň), ale jako hodnoty tří různých prvků.

4. Upřesnit typy vložených hodnot prvků formou hodnot atributů převzatých z řízených slovníků.

Pro udržení konzistence dat a možnosti jejich dalšího bezeztrátového automatického zpracování hrají klíčovou roli slovníky možných hodnot atributů. Jejich přítomnost zajistí možnost prověření správnosti dat – pokud nastane případ, kdy je jako hodnota atributu zadán výraz neuvedený ve slovníku, je toto možné zjistit automaticky a následně zajistit nápravu.

Konkrétní oblastí, na níž lze názorně doložit vhodnost použití hodnot atributů k zachycení upřesňujících charakteristik vlastních hodnot prvků, je uvádění názvů časopisů. Díky hodnotě atributu lze odlišit úplný název časopisu od názvu zkráceného, navíc je možné upřesnit použitý způsob zkracování¹⁸⁰. Může se jednat např. o zkracování v souladu s CASSI¹⁸¹, se zkracováním užívaným v systému ISSN (podle normy ISO 4¹⁸²), v databázi MEDLINE (resp. systému PubMed), v citačních rejstřících v rámci systému Web of Science či o specifický způsob vyjádření názvu v podobě identifikátoru CODEN.

Konkrétní způsob, jak tato data mohou být zakomponována do vytvářeného XML formátu, je znázorněn na příkladu časopisu *Acta Biochimica Polonica*:

```
<title>Acta Biochimica Polonica</title>  
<title convention="cassi">Acta Biochim. Pol.</title>  
<title convention="issn">Acta Biochim. Pol.</title>
```

¹⁸⁰ K dekódování/rozšifrování jednotlivých zkratk je možné využít online nástroj JAbbr vyvinutý na Cornellově univerzitě (*Cornell University*) [297].

¹⁸¹ CASSI je zkratka zdroje *Chemical Abstracts Service Source Index*. Nově je k dispozici online nástroj pro vyhledávání názvů časopisů exerpovaných službou Chemical Abstracts Service [298].

¹⁸² ISO 4:1997. *Information and documentation – rules for the abbreviation of title words and titles of publications*. Geneva : International Organization for Standardization, 1997.

<title convention="medline">Acta Biochim Pol</title>

<title convention="coden">ABPLAF</title>

<title convention="wos">ACTA BIOCHIM POL</title>

Podobně lze pro uvádění dat (popř. i časů) doporučit využití normy ISO 8601¹⁸³ [299].

5. Zachytit hierarchii citačních dat a využít odkazů k propojení jednotlivých hierarchických úrovní.

V citační praxi jsou často citovány dokumenty (nebo části dokumentů), které tvoří integrální součást jiného (zdrojového) dokumentu. Jako typické příklady mohou posloužit příspěvky v konferenčních sbornících, kapitoly v monografiích či zejména články v časopisech.

Proto se jeví jako vhodné rozčlenit v XML tato data tak, aby byly jednotlivé hierarchické úrovně (např. konkrétní stránka časopiseckého článku – časopisecký článek – číslo časopisu – ročník časopisu – časopis jako celek) uvedeny odděleně a propojeny prostřednictvím odkazů.

Díky tomu je možné jinak se opakující data uvést pouze jednou (např. údaje o časopisu jako celku) a následně na ně pouze odkazovat.

6. Využít existující typologii dokumentů nebo vytvořit novou typologii dokumentů a podle potřeby ji namapovat na již existující typologie dokumentů.

Vzhledem ke skutečnosti, že existuje řada různých typologií dokumentů, nelze jednoznačně doporučit konkrétní typologii, která by byla zvláště vhodná pro oblast citování. Lze proto doporučit dvě cesty. Jednou z nich je využití existující typologie dokumentů, druhou potom vytvoření vlastní typologie dokumentů. V obou případech může následně v praxi vzniknout potřeba namapování typologie na jinou existující typologii.

Z existujících typologií lze za vhodné pokládat zejména typologie dokumentů, které jsou volně dostupné, nejlépe na jednom místě na internetu, jsou postupně rozšiřované a rovněž doplněné o komentář týkající se použití typů dokumentů v praxi.

¹⁸³ ISO 8601:2004. *Data elements and interchange formats – Information interchange – Representation of dates and times*. Geneva : International Organization for Standardization, 2004.

V případě využití existující typologie dokumentů lze doporučit upřesnění použité typologie formou hodnoty atributu.

7. Umožnit využití principu rekurze při práci s daty.

Je rovněž potřeba zdůraznit využití rekurze (*recursion*), tj. jednoho z klíčových principů programování, resp. počítačové vědy, kdy řešení problému záleží na řešeních menších instancí téhož problému¹⁸⁴. Jinými slovy se jedná o proces definice či vyjádření funkce, procedury, jazykové konstrukce nebo řešení problému na základě sebe samé či samého [45].

Z programátorského hlediska se jedná o velmi výhodný princip, neboť díky rekurzi lze programování podstatně zefektivnit a zrychlit.

8.3 Návrh experimentálního formátu

8.3.1 Návrhová kritéria

Následující návrhová kritéria vycházejí z výše uvedených doporučených postupů. Jedná se o jedenáct kritérií:

1. Primárním cílem formátu bude zachytit údaje o dokumentech pro účely citování.
2. Formát bude využívat co nejnižší počet prvků (elementů) a atributů.
3. Prvky i atributy budou nazvány tak, aby byly pro člověka snadno srozumitelné.
4. Formát bude zkonstruován tak, aby byl použitelný pro širokou škálu aplikací (automatizovaných informačních systémů).
5. Formát bude možné využívat jak samostatně, tak v případech, kdy bude použit jiný formát a navržený formát bude sloužit pouze pro zachycení citačních dat.
6. Formát bude vytvořen v jazyce XML, a tedy bude strojově srozumitelný.
7. Přítomnost většiny prvků a atributů bude nepovinná.
8. Formát bude založen na využití slovníků hodnot atributů a vybraných prvků.
9. Formát bude umožňovat maximální možnou strukturaci dat.

¹⁸⁴ V tomto směru lze spatřovat analogii s fraktály.

10. Podstatnou roli bude hrát hierarchie dat a vzájemné odkazy.

11. Formát bude umožňovat využití principu rekurze při práci s daty v něm uloženými.

8.3.2 Metodika přípravy

Nejprve byla stanovena rámcová architektura XML formátu. Klíčovými východisky pro jeho architekturu se staly zásady uvedené v rámci doporučených postupů, návrhová kritéria a poznatky získané na základě dat (záznamů) z různých citačních stylů, resp. z dokumentů, v nichž jsou popsány.

První dvě východiska byla již blíže popsána v kap. 8.2 a 8.3. Pokud se týká třetího východiska, cílem bylo získat alespoň částečně reprezentativní vzorek typů záznamů, které je potřeba ve vědecké praxi citovat. Za účelem získání tohoto vzorku byli vybráni tři reprezentanti různě zaměřených citačních stylů – mezinárodní normy ISO 690 a ISO 690-2 (doplněné o české překlady těchto norem, tj. ČSN ISO 690 a ČSN ISO 690-2) a styly ACS a APA (v obou případech ve dvou nejnovějších vydáních). Mezinárodní normy jsou obecné, nejsou tedy určeny pro konkrétní tematickou oblast. Naopak styly ACS a APA se vyznačují orientací na oblasti chemie a psychologie.

Z těchto zdrojů byly excerpovány záznamy dokumentů. Excerpovaná data byla zachycena formou tabulek v textovém editoru OpenOffice.org Writer. Původní záměr vytvářet tabulky v tabulkovém procesoru OpenOffice.org Calc nebyl realizován. V takovém případě by bylo nutné do jedné buňky zapisovat dlouhé řetězce (celé záznamy), což by nebylo dostatečně přehledné. Naopak tabulky v textovém editoru v tomto směru vykazovaly vynikající vlastnosti.

Jeden záznam byl vždy zachycen do jedné tabulky, v níž byly v levé části uvedeny popisky dat a v pravé části vlastní data. Všechny tabulky měly stejnou podobu, byly složeny ze šesti řádků a dvou sloupců. V tabulkách bylo možné uvést identifikátory záznamů, znění konkrétního příkladu, uvedení typu dokumentu, poznámky, zdroje a stránky, z níž byl záznam excerpován. Příkladem je tab. 50.

Tab. 50: Příklad tabulky s excerpovanými citačními daty

Identifikátor	ACS3-291-1
Příklad	Carus, R. A.; Susha, A.; Caruso, F. Multilayered Titania, Silica, and Laponite Nanoparticle Coatings on Polystyrene Colloidal Templates and Resulting Inorganic Hollow Spheres. <i>Chem. Mater.</i> 2001 , <i>13</i> , 400–409.
Typ dokumentu	Periodicals
Poznamka	
Zdroj	The ACS Style Guide, 3rd ed.
Stranka	291

Pokud se týká identifikátorů, byly složeny ze tří částí. První část zachycovala kód zdrojového dokumentu (např. kód „ACS3“ zastupoval třetí vydání publikačního stylu ACS), následovala stránka, z níž byl záznam excerpován (v případě uvedení začátku záznamu na jedné stránce a konce na stránce druhé byla volena první stránka), a pořadí na stránce. Tyto tři údaje byly vždy oddělovány spojovníky.

Vlastní příklad (záznam) byl pak vždy uveden přesně v podobě, v níž se vyskytoval ve zdroji (tj. včetně případného typografického zvýraznění jednotlivých částí záznamu).

Typ dokumentu byl vždy převzat z daného zdroje, jehož zkrácený záznam byl také zahrnut do tabulky. Zařazena byla také stránka, z níž byl záznam excerpován. Tyto dva údaje byly zahrnuty z důvodu možnosti automatizovaného provedení kontroly správnosti záznamů.

Celkem bylo výše uvedeným způsobem zpracováno několik set tabulek, z nichž byla následně část převedena do první pracovní verze XML formátu. Pro přípravu souborů byl využit XML editor oXygen. Při zpracování byly získány cenné dílčí poznatky, které byly využity při první aplikaci formátu (viz kap. 8.3.3).

Ve formátu je jako kořenový prvek použit prvek „reference“¹⁸⁵, jemu jsou podřazeny prvky „document“ (jeden prvek nebo více prvků) složené ze dvou částí – prvku „head“ a prvku „body“. Prvek „head“ je rezervován pro budoucí užití pro uvedení „metadat“ týkajících se dat uvedených v prvku „body“ (např. úplného znění

¹⁸⁵ Názvy prvků jsou záměrně uváděny v anglickém jazyce, aby bylo možné s nimi pracovat nejenom v českém jazykovém prostředí.

záznamu z původního zdroje, data a času získání záznamu, poslední modifikace apod.). V prvku „body“ jsou uvedena vlastní citační data. Jednotlivé prvky „document“ jsou mezi sebou propojeny pomocí prvku „source“ s atributem „document-id“. Jako hodnota atributu jsou uvedeny identifikátory přidělené prvkům „document“, na něž prvek „source“ odkazuje. Rozčlenění dat do jednotlivých prvků „document“ je založeno na vztahu mezi citovaným dokumentem a jeho dílčími či naopak nadřazenými částmi. Je-li např. citována stránka z kapitoly knihy, která byla vydána v rámci určité edice, budou jednotlivé prvky „document“ zrcadlit tyto vztahy. V prvním prvku „document“ budou uvedeny údaje o stránce, ve druhém o kapitole, ve třetím o knize, ve čtvrtém o svazku v edici a v pátém o edici. Tato forma uvádění údajů velmi dobře umožňuje využití principu rekurze při zpracování dat.

Ve formátu je využito kódování UTF-8, zahrnut byl pouze nízký počet prvků a atributů. Vzhledem k tomu, že formát je zatím k dispozici v pracovní verzi, není účelné uvádět konečný počet ani úplný výčet jednotlivých prvků a atributů. V kap. 8.3.3 je proto uveden počet prvků a atributů a jejich výčet, který se vztahuje pouze k dané aplikaci formátu. Podobně je tomu v případě práce s řízenými slovníky a s konkrétní typologií dokumentů.

V pracovní verzi formátu je zatím použita vlastní typologie dokumentů, která vychází ze skutečnosti, že formát využívá hierarchické segmentace údajů, kdy se jako samostatný záznam objevuje např. článek, dále číslo časopisu, jehož je součástí, ročník a časopis samotný. Typologie v sobě proto musí mít zahrnut nejenom typ dokumentu článek, resp. „article“, ale také číslo, tj. „issue“, ročník, tj. „volume“ či celý časopis, tj. „journal“.

Pokud se týká principu strukturace dat, většina dat byla rozdělena na co nejmenší části. Jako názorný příklad může posloužit způsob zachycení autorských údajů (podrobnější popis je zařazen u příkladu v kap. 8.3.3):

```
<creators>
<author>
  <first-name type="initial">D.</first-name>
  <middle-name type="initial">W.</middle-name>
  <surname>Rogers</surname>
```

</author>
</creators>

8.3.3 Praktické využití formátu na příkladu citačních dat z barevných knih

Za účelem praktického ověření vhodnosti navrženého experimentálního formátu a získání podnětů pro jeho další úpravy a rozvoj byla zpracována část citačních dat ze tří barevných knih IUPAC – ze Zlaté, Červené a Zelené knihy. Konkrétně se jednalo o zpracování záznamů článků z časopisu *Pure and Applied Chemistry*, který je rovněž vydáván organizací IUPAC.

Při zpracování byla použita následující metodika:

Nejprve byla všechna citační data (záznamy v seznamech použitých zdrojů) excerpována ze zdrojových dokumentů a předzpracována formou jednotlivých tabulek. Jednalo se o stejný typ tabulek zmíněný výše v souvislosti se zpracováním citačních dat z publikačních manuálů APA, ACS a z citačních norem ISO 690 a ISO 690-2. Na rozdíl od těchto případů nebyla u barevných knih vyplňována hodnota typu dokumentu. Příklad tabulky je uveden jako tab. 51.

Tab. 51: Příklad tabulky s excerpovanými citačními daty

Identifikátor	GRE2-135-1
Příklad	Manual of Symbols and Terminology for Physicochemical Quantities and Units. 1st ed., McGlashan, M.L., <i>Pure Appl. Chem.</i> 21 (1970) 1–38.
Typ dokumentu	
Poznamka	
Zdroj	Green Book, 2nd ed.
Stranka	135

Každému excerpovanému záznamu byl přidělen identifikátor složený ze tří částí, z nichž první tvořil prefix přidělený pro účely dalšího zpracování dané barevné knihy, druhou číslo stránky, na níž se daný záznam nacházel, a třetí pořadové číslo na dané stránce. Jednotlivé části od sebe byly pro přehlednost a snazší vizuální kontrolu správnosti identifikátoru odděleny spojovníky.

V uvedených datech se tedy vyskytly celkem tři prefixy – GOLX, GRE2 a RED3. První tři znaky odkazují na slangový název dané barevné knihy (GOL – Gold Book, GRE – Green Book a RED – Red Book), poslední znak z prefixu potom upřesňuje vydání (tj. XML vydání, 2. vydání a 3. vydání).

Číslo stránky pak bylo aplikováno pouze v případě Červené a Zelené knihy, které jsou online k dispozici v podobě souborů ve formátu PDF (viz výše). V případě Zlaté knihy bylo číslo stránky nahrazeno pořadovým číslem skupiny, do nichž byly záznamy uspořádány (všechny zpracovávané záznamy byly v tomto případě umístěny na jedné webové stránce).

Pořadové číslo na stránce (resp. v případě Zlaté knihy ve skupině) bylo zvoleno vždy tak, že prvnímu záznamu bylo přiřazeno pořadové číslo 1, druhému 2 atd. (tj. pořadová čísla byla přiřazována vzestupně).

Po zpracování všech záznamů ze tří zmíněných barevných knih byly z těchto záznamů vybrány záznamy článků publikovaných v časopise Pure and Applied Chemistry. Celkem bylo tímto způsobem vybráno 199 záznamů, z toho ze Zlaté knihy pocházelo 131 záznamů, ze Zelené knihy 46 záznamů a z Červené knihy zbylých 22 záznamů.

V případě, kdy ve zdrojovém dokumentu bylo v rámci jedné položky (s citačními daty uvedenými v jedné tabulce) uvedeno více záznamů, bylo pořadové číslo doplněno o velká písmena podle abecedy. Příkladem takového identifikátoru je identifikátor „RED3-45-1C“ (3. vydání Červené knihy, strana 45, třetí záznam v první položce).

Vlastní zpracování citačních dat v XML souborech probíhalo tak, že nejprve byl vždy vyhledán úplný text daného článku v online archivu časopisu [300]. Ke zpracování byla využita jak data z webové stránky věnované danému článku, tak z vlastního úplného textu (ve formátu PDF).

Data z jednotlivých záznamů byla vkládána do XML souborů, které byly vždy nazvány tak, aby byl jejich název shodný s identifikátorem přiděleným záznamu (tj. např. „GRE2-135-1.xml“).

Celkem bylo použito šestnáct prvků:

author

body

contributor	place
creators	position
document	reference
first-name	source
head	surname
identifier	time
middle-name	title

Přehled použitých atributů těchto prvků a hodnot atributů podává tab. 52 (do tabulky jsou zahrnuty prvky, u nichž je použit alespoň jeden atribut).

Tab. 52: Použité prvky a atributy se svými hodnotami

Prvek	Atribut	Hodnota atributu
author	type	group
author	subtype	institution
contributor	type	edition
document	id	proměnlivá, vždy konkrétní identifikátor
first-name	type	initial
identifier	type	doi, url
middle-name	type	initial
place	type	city, country, state
	subtype	meeting
position	type	page
	subtype	start, end
source	document-id	proměnlivá, vždy konkrétní identifikátor
time	subtype	meeting
title	type	short, main, extended
	convention	cassi, iupac, meeting

Konkrétní podobu zpracovaných XML souborů budeme ilustrovat na příkladu souboru s identifikátorem GOLX-3-72:¹⁸⁶

```
<?xml version="1.0" encoding="UTF-8"?>
<reference>
  <document id="GOLX-3-72">
```

¹⁸⁶ Uvedený příklad zatím nepracuje se jmennými prostory. Ty budou zavedeny až ve finální verzi formátu.

```

<head/>
<body>
  <type>journal-article</type>
  <title>Nomenclature, symbols, units and their usage in spectrochemical analysis -
    XII. Terms related to electrothermal atomization</title>
  <creators>
    <author type="group" subtype="institution">International Union of Pure and
      Applied Chemistry, Analytical Chemistry Division, Commission on
      Spectrochemical and Other Optical Procedures for Analysis</author>
    <contributor type="edition">
      <first-name type="initial">T.</first-name>
      <middle-name type="initial">A.</middle-name>
      <middle-name type="initial">M.</middle-name>
      <surname>Ure</surname>
    </contributor>
    <contributor type="edition">
      <first-name type="initial">L.</first-name>
      <middle-name type="initial">R.</middle-name>
      <middle-name type="initial">P.</middle-name>
      <surname>Butler</surname>
    </contributor>
    <contributor type="edition">
      <first-name type="initial">B.</first-name>
      <middle-name type="initial">V.</middle-name>
      <surname>L'vov</surname>
    </contributor>
    <contributor type="edition">
      <first-name type="initial">I.</first-name>
      <middle-name type="initial">Rubeska</middle-name>
      <surname></surname>
    </contributor>
    <contributor type="edition">
      <first-name type="initial">R.</first-name>
      <middle-name type="initial">Sturgeon</middle-name>
      <surname></surname>
    </contributor>
  </creators>
  <position type="page" subtype="start">253</position>

```

```

    <position type="page" subtype="end">259</position>
    <identifier type="doi">doi:10.1351/pac199264020253</identifier>
    <identifier type="url">http://iupac.org/publications/pac/64/2/0253/</identifier>
    <source document-id="GOLX-3-72a"/>
  </body>
</document>

<document id="GOLX-3-72a">
  <head/>
  <body>
    <type>issue</type>
    <title type="short">2</title>
    <time>1992</time>
    <source document-id="GOLX-3-72b"/>
  </body>
</document>

<document id="GOLX-3-72b">
  <head/>
  <body>
    <type>volume</type>
    <title type="short">64</title>
    <time>1992</time>
    <source document-id="GOLX-3-72c"/>
  </body>
</document>

<document id="GOLX-3-72c">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>
</reference>

```

Soubor začíná deklarací, která určuje použitou verzi jazyka XML (1.0) a také kódování (UTF-8). Vlastní data jsou pak uložena v prvku „reference“. Jemu

jsou podřízeny prvky „document“, v nichž jsou uložena dílčí údaje vztahující se k jednotlivým hierarchickým úrovním dat. Každý prvek dokument obsahuje prvek „head“ (v tomto případě zatím prázdný) a prvek „body“.

Prvku „body“ jsou již podřízeny další prvky – ve výše uvedeném případě se jedná o sedm prvků: „type“, „title“, „creators“, „position“, „identifier“, „source“ a „time“.

Prostřednictvím hodnoty prvku „type“ je upřesněn typ citovaného dokumentu, resp. hierarchická úroveň dat od časopiseckého článku (hodnota „journal-article“) přes číslo (hodnota „issue“) a ročník (hodnota „volume“), v nichž byl článek zveřejněn, až po celý časopis (hodnota „journal“).

Prvek „title“ je využíván k zachycení názvů – v prvním případě se jedná o název článku, v dalších případech o názvy čísla, ročníku a časopisu. Zatímco v případě názvu článku a názvu časopisu je typ vyplňovaných údajů zřejmý, názvy čísla a ročníku si zaslouží bližší vysvětlení. V tomto případě jsou totiž jako hodnoty prvku vyplněna číselná označení. U prvku je v těchto případech doplněn atribut „type“ s hodnotou „short“, neboť je vhodné počítat i s možností, kdy by nebylo potřeba zachytit jenom zmíněná číselná označení, ale také odpovídající slovní vyjádření, např. formou zkratk „no.“, „vol.“ apod. Zachycení názvu časopisu také názorně ukazuje možnost vícenásobného využití prvku v rámci jednoho prvku „document“ a také možnost přesného určení způsobu zkrácení názvu pomocí atributu „convention“. Ve výše uvedeném případě se jedná o využití konvence zkracování podle CASSI, doplňkově je pak zařazena zkratka užívaná v organizaci IUPAC.

Dále je použit prvek „creators“, jemuž jsou podřízeny prvky „author“ a „contributor“. Zatímco první uvedený prvek slouží k uvádění primární autorské odpovědnosti k citovanému dokumentu, druhý prvek je určen pro údaje o sekundárních autorech. V našem případě náleží primární autorská odpovědnost korporaci (organizaci), proto jsou u prvku „author“ doplněny atribut „type“ s hodnotou „group“ a atribut „subtype“ s hodnotou „institution“. Jedná-li se o personální autorství, k uvedení konkrétních jmen jsou využity prvky „first-name“, „middle-name“ a „surname“. V případě, že nejsou známa úplná znění prvních

a středních jmen, jsou prvky „first-name“ a „middle-name“ doplněny o atribut „type“ s hodnotou „initial“.

Dalším prvkem je prvek „position“, který slouží k uvádění lokace citovaného dokumentu. V příkladu je tento prvek využit v kombinaci s atributem „type“ s hodnotou „page“ a s atributem „subtype“ s hodnotami „start“ a „end“. Díky tomu je možné přesně odlišit údaje o počáteční a koncové lokaci dokumentu.

Velmi důležitou úlohu hraje prvek „identifier“, v tomto konkrétním případě je prvek využit s atributem „type“ a hodnotami „url“ a „doi“. Právě díky zařazení identifikátoru DOI bylo možné snadno a přesně zjistit, že ve zpracovaném vzorku dat z barevných knih bylo více než jednou citováno celkem 35 článků (ve 34 případech byly dané články citovány dvakrát, v jednom případě třikrát).

Na využití identifikátorů je založen i celý formát – v našem případě je celému souboru a prvnímu výskytu prvku „document“ přiřazen identifikátor „GOLX-3-72“, pro identifikaci dalších výskytů tohoto prvku je využito téhož identifikátoru doplněného ovšem o malá písmena („a“, „b“ a „c“). K vzájemnému odkazování zdrojových dokumentů na jednotlivých hierarchických úrovních slouží prvek „source“ s atributem „document-id“, jehož hodnotou je vždy hodnota identifikátoru odkazovaného prvku „document“.¹⁸⁷

Při zpracování dat byla zjištěna řada poznatků, které budou následně využity při dalším vylepšení navrhovaného formátu. Jako užitečné vylepšení se např. jeví přidání hodnot „prefix“ a „suffix“ do slovníků atributů „type“ u prvku „surname“. Ve zpracovávaných záznamech se vyskytla řada případů, kdy mohlo být toto rozlišení úspěšně využito (viz tab. 53).

¹⁸⁷ V této experimentální aplikaci byl v jednom záznamu vždy uváděn pouze jeden identifikátor a jeho odvozeniny. Pro vlastní nasazení v praxi je samozřejmě výhodnější použití stejných identifikátorů v případě, že se vztahují k totožným datům (např. údajům o časopise jako celku).

Tab. 53: Příklady hodnot prvků „surname“

Identifikátor záznamu	Hodnota prvku „surname“
GOLX-3-112	de Bolster
GOLX-3-113	vam de Waterbeemd
GOLX-3-21	de Bruin
GOLX-3-27	den Boef
GOLX-3-84	Van Grieken
GOLX-3-84	de Bruin
GOLX-3-92	van der Linden
GRE2-135-6	Van Olphen
GRE2-135-8	Burwell, Jr.
GRE2-138-11	De Bièvre
GRE2-138-11	De Laeter
GRE2-138-14	Lide, Jr.
RED3-14-13	de Bolster

Např. místo podoby `<surname>de Bolster</surname>` by pak prvek mohl mít podobu `<surname type="prefix">de</surname> <surname> Bolster</surname>`. Obdobně lze řešit případ, kdy za příjmením následuje ještě sufix, tj. místo `<surname>Burwell, Jr.</surname>` lze pracovat s podobou `<surname>Burwell</surname> <surname type="suffix">Jr.</surname>`.

Dále byla identifikována vhodnost podrobnějšího rozčlenění prvku „author“ v případech, kdy je potřeba zachytit primární korporativní autorství. Převažující část zpracovávaných záznamů časopiseckých článků měla právě tento typ primárního autorství a navíc se v řadě případů jednalo o více než jednu organizaci, popř. její složku. V této první fázi zpracování byla potřebná data v daných záznamech uvedena jako hodnota prvku „author“ s tím, že v případech, kdy byly upřesněny organizační složky, byly jednotlivé dílčí údaje odděleny čárkou a mezerou. Např. v souboru s identifikátorem GOLX-3-12 měla hodnota prvku podobu `<author type="group" subtype="institution">International Union of Pure and Applied Chemistry, Analytical Chemistry Division, Commission on Microchemical Techniques and Trace Analysis</author>`.

8.3.4 Vizualizace citačních vztahů

Na základě XML souborů se zpracovanými záznamy článků z časopisu Pure and Applied Chemistry, které byly citovány ve Zlaté, Červené a Zelené knize, byla zpracována experimentální vizualizace citačních vztahů (připravil ji Ing. Miloslav Nič, Ph.D.). Konkrétně se jednalo o vizualizaci, která zobrazuje dokumenty sdílející alespoň jednoho (personálního) autora či přispěvatele.¹⁸⁸

Při vizualizaci byly využity jazyky Python a XSLT. Jejich prostřednictvím byl vytvořen soubor ve formátu DOT, tj. interního formátu softwarové aplikace Graphviz (viz obr. 82). Pomocí této aplikace byla vytvořena zmíněná vizualizace citačních vztahů ve formátu SVG¹⁸⁹ a následně PNG¹⁹⁰ v editoru Inkscape určeném pro vektorovou grafiku.

```
digraph G {
  "GRE2-138-11" -> "GRE2-138-10";
  "RED3-52-2" -> "RED3-14-3";
  "RED3-14-12B" -> "GOLX-3-102";
  "RED3-14-12B" -> "GOLX-3-41";
  "RED3-14-12B" -> "GOLX-3-130";
  "RED3-14-12B" -> "GOLX-3-102";
  "RED3-14-12B" -> "GOLX-3-128";
  "RED3-14-12B" -> "GOLX-3-125";
  "RED3-14-12B" -> "GOLX-3-130";
  "RED3-14-12B" -> "GOLX-3-102";
  "RED3-14-12B" -> "GOLX-3-122";
  "RED3-14-12B" -> "GOLX-3-124";
  "RED3-14-12B" -> "GOLX-3-128";
  "RED3-14-12B" -> "GOLX-3-102";
  "GOLX-3-17" -> "GRE2-135-11";
  "GOLX-3-17" -> "GOLX-3-104";
  "RED3-13-13A" -> "GOLX-3-93";
  "RED3-13-13A" -> "RED3-14-7D";
  "RED3-13-13A" -> "GOLX-3-103";
  "RED3-13-13A" -> "RED3-45-1C";
  "RED3-13-13A" -> "GOLX-3-58";
  "RED3-13-13A" -> "GOLX-3-47";
  "GOLX-3-119" -> "GOLX-3-120";
  "GOLX-3-119" -> "GOLX-3-81";
  "GOLX-3-46" -> "GRE2-137-24";
  "GRE2-135-7" -> "GOLX-3-30";
  "GRE2-135-7" -> "GOLX-3-30";
  "RED3-13-14" -> "RED3-110-2A";
  "RED3-13-14" -> "RED3-110-2A";
}
```

Obr. 82: Ukázka souboru ve formátu DOT

Ukázku výsledných vizualizovaných vztahů přinášejí obr. 83 a 84. Na obr. 83 je zachycena pouze malá část vizualizovaných vztahů, konkrétně je vybrána část

¹⁸⁸ Za dokumenty sdílející alespoň jednoho autora či přispěvatele byly pokládány ty, u nichž byly autorské údaje uvedeny ve zcela shodné podobě. Nejsou tedy zahrnuty případy, kdy by například v jednom dokumentu měl autor uvedeno úplné znění prvního jména a ve druhém pouze jeho iniciálu.

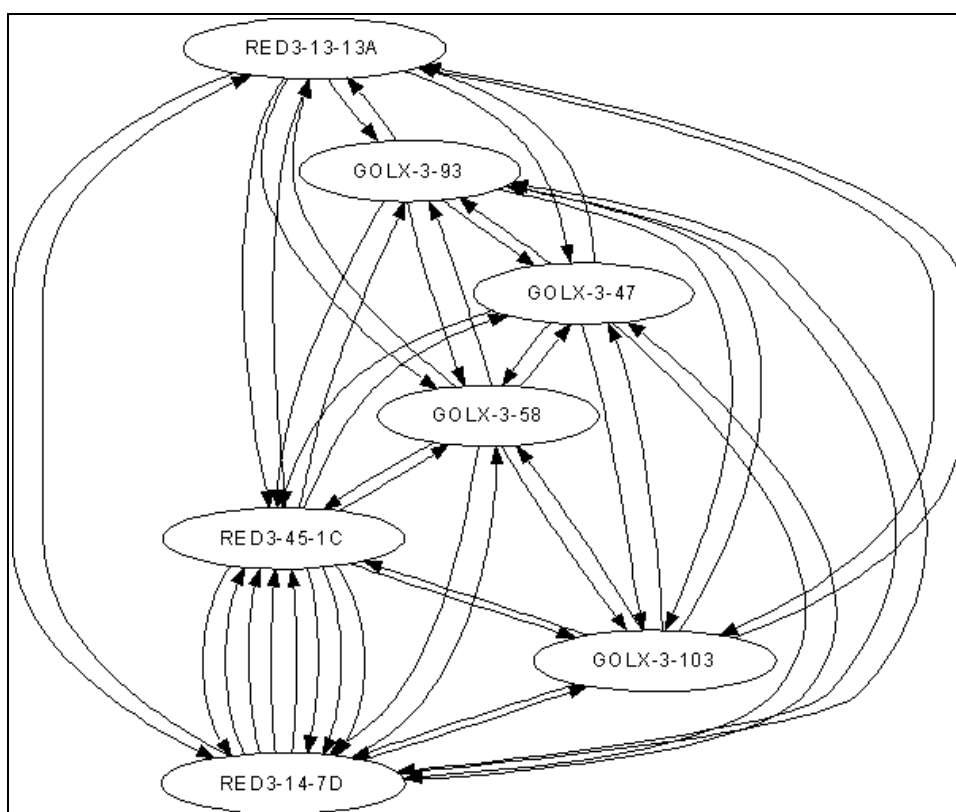
¹⁸⁹ SVG je zkratka výrazu *Scalable Vector Graphics* (škálovatelná vektorová grafika).

¹⁹⁰ PNG je zkratka výrazu *Portable Network Graphics* (přenositelná síťová grafika).

znázorňující vztahy mezi sedmi XML soubory. Z nich se ve čtyřech případech jedná o záznamy dokumentů ze Zlaté knihy a ve třech případech o záznamy z Červené knihy. Pro názornost jsou příslušné XML soubory k této práci připojeny jako příl. č. 15-21.

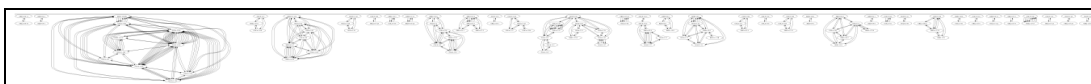
Dokumenty ve vizualizaci představují uzly grafu, vztahy mezi nimi pak hrany grafu. Vztah mezi dvěma dokumenty je obousměrný, přičemž tato obousměrnost je naznačena šipkami směřujícími k uzlům. Plyne to ze skutečnosti, že autor společný oběma dokumentům je uveden v obou těchto dokumentech zároveň.

V případě dvojice GOLX-3-47 a RED3-13-13A a dvojice RED3-14-7D a RED3-45-1C se jedná o totožné dokumenty, tedy o dokumenty, které sdílejí všechny autory.



Obr. 83: Vizualizace části vztahů mezi dokumenty

Obr. 84 je makropohledem na všechny identifikované vztahy mezi dokumenty. Byla tedy zjištěna existence většího počtu samostatných shluků dokumentů sdílejících více autorů.



Obr. 84: Vizualizace všech identifikovaných vztahů mezi dokumenty

Ve výše uvedeném případě byly tedy vizualizovány vztahy mezi dokumenty na základě společných autorů. Podobně budeme-li například mít u dat týkajících se monografií (či obecně knih) uvedeno místo vydání (město, zemi, popř. další upřesnění, např. jednotlivý stát v USA), je možné pro danou množinu dat reprezentujících monografie vytvořit vizualizovaný výstup, díky kterému bude možné zjistit koncentraci vydávaných monografií z dané množiny na různých místech na světě.

Je tedy možné shrnout, že na základě dat v XML formátu je také významně usnadněna možnost vizualizace vztahů mezi jednotlivými prvky.

8.4 Dílčí závěry

Pro zahrnutí citačních dat do XML formátů bylo formulováno celkem sedm klíčových zásad:

1. Využívat implicitní kódování XML, tj. UTF-8.
2. Pracovat s relativně nízkým počtem prvků i atributů.
3. Co nejvíce strukturovat citační data, aby bylo následně možné efektivně provádět jejich potřebné transformace.
4. Upřesnit typy vložených hodnot prvků formou hodnot atributů převzatých z řízených slovníků.
5. Zachytit hierarchii citačních dat a využít odkazů k propojení jednotlivých hierarchických úrovní.
6. Využít existující typologii dokumentů nebo vytvořit novou typologii dokumentů a podle potřeby ji namapovat na již existující typologie dokumentů.
7. Umožnit využití principu rekurze při práci s daty.

Na základě těchto zásad bylo následně přistoupeno k formulaci jedenácti návrhových kritérií pro přípravu experimentálního formátu.

V první fázi byla provedena excerptce dat z publikačních manuálů APA a ACS a z norem ISO 690 a 690-2. Část těchto dat byla následně ručně převedena do první verze vytvářeného XML formátu.

Poznatky získané z předchozího kroku byly prakticky aplikovány při zpracování citačních dat ze tří barevných knih IUPAC. Tato aplikace rovněž přispěla ke zpřesnění formátu. Jako užitečné se ukázalo přidání hodnot „prefix“ a „suffix“ do slovníků atributů „type“ u prvku „surname“ a dále podrobnější rozčlenění obsahu prvku „author“ v případech, kdy je potřeba zachytit primární korporativní autorství.

Následná vizualizace vztahů mezi citačními daty na příkladu společných autorů citovaných dokumentů ukázala, že s daty v navrženém XML formátu je možné efektivně dále pracovat a provádět jejich transformace.

9. Závěr

Problematika citování představuje bezesporu jednu z důležitých oblastí informační vědy i knihovnicko-informační praxe, přesto nelze než konstatovat, že míra standardizace citování zatím není dostatečná. Skutečnost, že v oblasti vědecké literatury nejsou uplatňovány jednotné postupy pro tvorbu záznamů citovaných dokumentů, názorně doložil provedený detailní průzkum citačních stylů používaných v padesáti časopisech s nejvyšším impakt faktorem. V záznamech vytvořených na základě různých citačních stylů je uváděna různá interpunkce, liší se také například pořadí jednotlivých údajů. Tyto skutečnosti velmi komplikují automatizované zpracování citačních dat.

Jedním z možných řešení je automatická extrakce citační dat. Ta je však často zatížena nepřesnostmi a chybami. Dalším možným a moderním přístupem je vytvoření XML formátu pro zachycení citačních dat. Mezi hlavní výhody XML patří existence přesné specifikace jazyka, možnost vytvořit schémata využitelná ke kontrole správnosti zadaných dat a možnost snadno provádět transformace dat. Proto byl v předkládané práci zvolen právě přístup založený na XML.

Klíčovým předpokladem pro jeho aplikaci je provedení analýzy způsobů zahrnutí citačních dat do vybraných XML formátů z oblasti STM a analýzy existujících citačních formátů a jejich uplatnění v praxi. Bylo zjištěno, že v některých XML formátech je zahrnutí citačních dat zpracováno velmi podrobně. Jednalo se zejména o formáty DocBook a TEI. Formát EPUB byl naopak identifikován jako příklad formátu, který nemá tuto oblast detailně rozpracovanou. V poslední době se také rozvíjejí projekty zaměřené na strojům srozumitelná citační data, zejména projekt jazyka pro citační styly CSL a projekty ontologií BIBO a CiTO. Pokud se týká citačních formátů, často jsou v praxi implementovány formáty RIS, EndNote, RefWorks či BibTeX. V prvních třech případech se jedná o formáty určené primárně pro import dat do citačních manažerů Reference Manager, EndNote a RefWorks; formát BibTeX byl původně vyvinut pro potřeby typografického systému LaTeX. Objevují se také formáty v jazyce XML, např. MARCXML, MODS či NLM.

Průzkum výskytu citačních formátů ve čtrnácti vybraných informačních systémech ukázal, že výstupy v citačních formátech v jazyce XML jsou zatím spíše výjimkou než pravidlem. Přesto lze konstatovat, že jazyk XML se postupně prosazuje jako jazyk strojově srozumitelných citačních formátů. Řada citačních formátů, které vznikly v době před rozšířením jazyka XML, byla aktuálně doplněna o svou verzi právě v tomto jazyce. Nově vznikají citační formáty v jazyce XML, které nejsou pouze jednoduchou obdobou dřívějších formátů nezaložených na XML, ale výrazněji využívají možností XML.

Navíc jsou citační formáty v XML již postupně implementovány i do citačních manažerů. Jako příklad mohou posloužit nástroje Zotero či Connotea dostupné prostřednictvím internetového prohlížeče. Zotero nabízí tři XML formáty pro import a export dat, Connotea potom čtyři XML formáty pro export dat. Zvláště u citačních formátů nevyužívajících XML se jako velmi podstatný nedostatek ukázalo nevyhovující kódování. Názorným příkladem je formát RIS, v jehož specifikaci je předepsána rozšířená znaková sada IBM. Ta však současným potřebám již nevyhovuje, a proto je v praxi (navzdory specifikaci) často implementováno kódování UTF-8.

Problematický je také nejednotný způsob uvádění vlastních citačních dat, kdy jsou například jména autorů uváděna někdy v invertované a někdy v přímé podobě. Podobně ne vždy jsou rozlišovány úplné názvy časopisů od názvů zkrácených.

Značná různorodost způsobů uvádění citačních dat byla zjištěna i v barevných knihách vydávaných organizací IUPAC. Je zřejmé, že podoba citací a záznamů v barevných knihách se neřídí jednotnými citačními instrukcemi či pokyny, ačkoliv bychom mohli předpokládat, že tomu ani jinak nemůže být u fundamentální série publikací vydávaných renomovanou vědeckou společností.

Podobně různorodé jsou také typologie dokumentů, s nimiž pracují citační formáty, styly a ontologie, potažmo typologie, které jsou využívány v oblasti knihovnictví. Lze je pokládat za velmi pragmatické, navíc ne ve všech případech jsou dostupné prostřednictvím internetu a doplněné o komentář týkající se aplikace typů dokumentů v praxi.

Neuplatňování jednotných standardů komplikuje možnost automatického zpracování citačních dat a výrazně omezuje využitelnost těchto dat pro další účely.

Na základě dílčích zjištění prezentovaných v kapitolách 2 až 7 jsem formulovala těchto sedm klíčových zásad pro zahrnutí citačních dat do XML formátů:

1. Využívat implicitní kódování XML, tj. UTF-8.
2. Pracovat s relativně nízkým počtem prvků i atributů.
3. Co nejvíce strukturovat citační data, aby bylo následně možné efektivně provádět jejich potřebné transformace.
4. Upřesnit typy vložených hodnot prvků formou hodnot atributů převzatých z řízených slovníků.
5. Zachytit hierarchii citačních dat a využít odkazů k propojení jednotlivých hierarchických úrovní.
6. Využít existující typologii dokumentů nebo vytvořit novou typologii dokumentů a podle potřeby ji namapovat na již existující typologie dokumentů.
7. Umožnit využití principu rekurze při práci s daty.

Na základě doporučených postupů byl vytvořen návrh experimentálního XML formátu pro citační data. Ten byl také využit pro zachycení části citačních dat ze Zlaté, Červené a Zelené knihy. Konkrétně bylo zpracováno celkem 199 záznamů článků z časopisu Pure and Applied Chemistry. Celkem bylo pro vyjádření citačních dat použito šestnáct prvků („author“, „body“, „contributor“, „creators“, „document“, „first-name“, „head“, „identifier“, „middle-name“, „place“, „position“, „reference“, „source“, „surname“, „time“ a „title“) a pět atributů („id“, „document-id“, „type“, „subtype“, „convention“). Vizualizace vztahů mezi citačními daty na příkladu společných autorů citovaných dokumentů doložila, že s daty v navrženém XML formátu je možné efektivně dále pracovat a provádět jejich transformace.

Problematika citování prostřednictvím XML samozřejmě není zdaleka předloženými výsledky uzavřena či vyčerpána. Pro další výzkum a praktickou aplikaci jeho výstupů se otevírají různé možnosti. Patří mezi ně specifikace obsahu prvku označeného jako „head“, příprava řízených slovníků hodnot atributů

i některých prvků a následné ověření jejich vhodnosti a praktické použitelnosti na vzorku dat. V tomto případě je možné využít již předzpracovaných záznamů z publikačních manuálů ACS a APA a z norem ISO 690 a 690-2 (popř. i z jejich českých překladů). Jako vhodný se jeví iterativní postup, tj. přibližování se výslednému řešení opakováním jednotlivých kroků a postupným zlepšováním jejich výstupů.

Seznam použitých pramenů a literatury

Poznámky:

Záznamy jsou seřazeny abecedně a jsou zpracovány v souladu s ČSN ISO 690, resp. ČSN ISO 690-2. V případech, kdy nebyl k dispozici název dokumentu, byl vytvořen umělý název a uveden v hranatých závorkách. Záznamy online verzí barevných knih jsou vytvořeny tak, aby reflektovaly primárně tyto své verze, tj. jejich součástí nejsou podrobné údaje týkající se původního tištěného vydání, na jehož základě byly online verze realizovány. Podrobnější údaje byly v případě potřeby zahrnuty do vlastního textu předložené práce.

1. *Extensible Markup Language (XML) 1.0* [online]. 5th ed. W3C, c2008, 26 November 2008 [cit. 2010-03-16]. Dostupné na World Wide Web: <<http://www.w3.org/TR/xml/>>.
2. *Extensible Markup Language (XML) 1.1* [online]. 2nd ed. W3C, 2006, 16 August 2006, edited in place 29 September 2006 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.w3.org/TR/xml11/>>.
3. JANSOVÁ, Linda; NIČ, Miloslav. Citování a citační praxe v digitálním prostředí. *ITlib : informačné technológie a knižnice*. 2009, roč. 13, č. 4, s. 13-20. Dostupné také na World Wide Web: <http://www.cvtisr.sk/itlib/itlib094/jansova_nic.htm>. ISSN 1335-793X.
4. SKOLKOVÁ, Linda. Managing Citations and References with Software Tools. In *15th International Seminar CASLIN 2008 : Knowledge Technologies nad (the Future of) Higher Education*. Brno : Masaryk University, 2008, s. 42-52. Prezentace dostupná na World Wide Web: <<http://alx.upol.cz/caslin2008/pom/skolkova.ppt>>. Videozáznam dostupný na World Wide Web: <<http://www.inflow.cz/caslin-2008-streamovane-video>>. ISBN 978-80-210-4614-6.
5. SKOLKOVÁ, Linda; NIČ, Miloslav. Institutional repositories and document citation. In *CASLIN 2009 : Institutional Online Repositories and Open Access : 16th International Seminar : Teplá Monastery, Czech Republic, 7-11 June 2009*. Pilsen : University of West Bohemia, 2009, s. 50-60. Prezentace dostupná na World Wide Web: <http://alx.upol.cz/caslin2009/prezentace/CASLIN_2009_Skolkova.ppt>. ISBN 978-80-7043-806-0.
6. *Ontology:bibo* [online]. [cit. 2010-02-03]. Dostupné na World Wide Web: <<http://bibotools.googlecode.com/svn/bibo-ontology/trunk/doc/index.html>>.
7. SHOTTON, David. *Ontology: cito* [online]. Version 1.6. 2010-03-26 [cit. 2010-05-15]. Dostupné na World Wide Web: <<http://imageweb.zoo.ox.ac.uk/pub/2009/citobase/cito-20100528-1.6/cito-content/owldoc/>>.
8. SHOTTON, David. *CiTO, the Citation Typing Ontology* [online]. [cit. 2010-06-10]. Článek nabídnutý k publikaci ve zvláštním čísle BioOntologies 2009 časopisu Journal of Biomedical Semantics. Dostupné na World Wide Web: <http://imageweb.zoo.ox.ac.uk/pub/2009/publications/Shotton_CiTO_preprint_and_suppl_files_for_JBMS_131109.pdf>.
9. KOSATA, Beda. *citace ve XML formátu* [elektronická pošta]. Message to: Linda Jansova. 6. 4. 2010 [cit. 2010-06-11]. Osobní komunikace.

10. *[DTD XML verze formátu EndNote pro EndNote Web]* [online]. Thomson Reuters, EndNote [cit. 2010-03-15]. Dostupné na World Wide Web: <<http://www.endnote.com/support/helpdocs/endnote.zip>>.
11. *EndNote* [počítačový program]. Version X3 (Bld. 4094). Thomson Reuters, c1988-2009 [cit. 2010-03-02]. Zkušební verze (k dispozici po registraci). Dostupné na World Wide Web: <<http://www.endnote.com/ndemo.asp>>.
12. *root/extension/branches/2.0/system.sql*. In *Zotero Issue Tracker* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research [cit. 2010-06-02]. Dostupné na World Wide Web: <<https://www.zotero.org/trac/browser/extension/branches/2.0/system.sql>>.
13. *Connotea: free online reference management for all researchers, clinicians and scientists* [online]. Nature Publishing Group, c2005-2010 [cit. 2010-05-01]. Connotea Code. Dostupné na World Wide Web: <<http://www.connotea.org/code#nam>>.
14. *CiteULike* [online]. Oversity [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.citeulike.org/>>.
15. *RIS Format Specifications* [online]. Thomson Reuters, This page was last modified on: February 14, 2001 [cit. 2010-03-25]. Reference Type Field Names. Dostupné na World Wide Web: <http://www.refman.com/support/risformat_reftypes.asp>.
16. *Web of Science Help* [online]. Thomson Reuters, c2009, This help page last modified 02/17/2009 [2010-05-01]. Document Types. Dostupné na World Wide Web: <http://images.isiknowledge.com/WOK45/help/WOS/hs_doctype.html>.
17. *RefWorks Tagged Format* [online]. RefWorks, [cit. 2010-03-30]. Dostupné na World Wide Web: <http://www.refworks.com/rwathens/help/RefWorks_Tagged_Format.htm>.
18. *EndNote X3 : User's Guide* [online]. Thomson Reuters, c2009 [cit. 2010-03-02]. 799 s. Určeno pro verzi softwaru pro operační systém Microsoft Windows. Dostupné na World Wide Web: <<http://www.endnote.com/support/helpdocs/EndNoteX3WinManual.pdf>>.
19. MARKEY, Nicolas. *Tame the BeaST : The B to X of BibTeX* [online]. Version 1.4. October 11, 2009 [cit. 2010-02-14]. 48 s. Dostupné na World Wide Web: <ftp://tug.ctan.org/pub/tex-archive/info/bibtex/tamethebeast/tt_en.pdf>.
20. *Registry for the OpenURL Framework - ANSI/NISO Z39.88-2004* [online]. NISO; OCLC [cit. 2010-05-15]. Matrix defining the KEV Format to represent a Book. Dostupné na World Wide Web: <<http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:book>>.
21. *Registry for the OpenURL Framework - ANSI/NISO Z39.88-2004* [online]. NISO; OCLC [cit. 2010-05-15]. Matrix defining the KEV Format to represent a journal publication. Dostupné na World Wide Web: <<http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:journal>>.
22. Library of Congress. Network Development and MARC Standards Office. *MARC 21 Format for Bibliographic Data* [online]. 1999 Edition. Update No. 1 (October 2001) through Update No. 11 (February 2010). Washington (DC, USA) : Library of Congress, Network Development and MARC Standards Office, February 2010 [cit. 2010-05-03]. 006 - Fixed-Length Data Elements-Additional Material Characteristics. Dostupné na World Wide Web: <<http://www.loc.gov/marc/bibliographic/bd006.html>>.
23. Library of Congress. *MODS : Metadata Object Description Schema Official Website* [online]. Washington (DC, USA) : Library of Congress, April 9, 2010 [cit. 2010-04-10]. Outline of Elements and Attributes in MODS Version 3.3. Dostupné na World Wide Web: <<http://www.loc.gov/standards/mods/mods-outline.html>>.

24. *PubMed Help* [online]. PubMed, 2005- , Last Update: June 7, 2010. [cit. 2010-06-08]. Publication Types. Dostupné na World Wide Web: <<http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helppubmed&part=pubmedhelp&rendertype=table&id=pubmedhelp.T42>>.
25. *Publication Manual of the American Psychological Association*. 5th ed. Washington (DC, USA) : American Psychological Association, c2001. xxviii, 439 s. ISBN 1-55798-791-2.
26. *Publication Manual of the American Psychological Association*. 6th ed. Washington (DC, USA) : American Psychological Association, c2010. xviii, 272 s. ISBN 978-1-4338-0561-5.
27. DODD, Janet S. (ed.). *The ACS style guide : a manual for authors and editors*. 2nd ed. Washington (DC, USA) : American Chemical Society, 1997. xii, 460 s. ISBN 0-8412-3462-0.
28. COGHILL, Anne M.; GARSON, Lorrin R. (ed.). *The ACS style guide : effective communication of scientific information*. 3rd ed. Washington (DC, USA) : American Chemical Society; Oxford (Velká Británie) : Oxford University Press, 2006. xiv, 430 s. ISBN 0-8412-7400-2.
29. ISO 690. *Documentation – Bibliographic references – Content, form, and structure*. 2nd ed. [Geneva] : International Organization for Standardization, 1987. iii, 11 s.
30. ISO 690-2. *Information and documentation – Bibliographic references – Part 2: Electronic documents or parts thereof*. 1st ed. [Geneva] : International Organization for Standardization, 1997. v, 18 s.
31. *CiTO, the Citation Typing Ontology: Vocabulary of Terms* [online]. [cit. 2010-02-02]. 6 s. Dostupné na World Wide Web: <http://imageweb.zoo.ox.ac.uk/pub/2009/citobase/CiTO_terms.doc>.
32. *Anglo-American Cataloging Rules : Second Edition : 1998 Revision*. Amendments 2001. Prepared under the direction of the Steering Committee for Revision of AACR. Ottawa : Canadian Library Association, 2001. Volné listy.
33. *RDA : Full Draft* [online]. American Library Association; Canadian Library Association; The Chartered Institute of Library and Information Professionals, c2008 [cit. 2010-06-15]. Dostupné na World Wide Web: <http://www.rdatoolkit.org/constituencyreviewfiles/constituency_review.zip>.
34. UDC Consortium. *UDC Consortium* [online]. The Hague : UDC Consortium, 2009 [cit. 2010-03-06]. UDC Summary. Dostupné na World Wide Web: <<http://udcc.org/udccsummary/php/index.php>>.
35. *Quantities, Units and Symbols in Physical Chemistry* [online]. Prepared for publication by Ian Mills, Tomislav Cvitaš, Klaus Humann, Kozo Kuchitsu. 2nd ed. London : Blackwell Science, c1993 [cit. 2009-09-03]. Dostupné na World Wide Web: <http://media.iupac.org/publications/books/gbook/green_book_2ed.pdf>. ISBN 0-632-03583-8.
36. *Nomenclature of Inorganic Chemistry* [online]. Issued by the Division of Chemical Nomenclature and Structure Representation in collaboration with the Division of Inorganic Chemistry. Prepared for publication by Neil G. Connelly, Ture Damhus, Richard M. Hartshorn, Alan T. Hutton. Cambridge (Velká Británie) : RSC, c2005 [cit. 2009-09-03]. xii, 366 s. Dostupné na World Wide Web: <http://media.iupac.org/publications/books/rbook/Red_Book_2005.pdf>.
37. *IUPAC Nomenclature of Organic Chemistry* [online]. Toronto (Kanada) : Advanced Chemistry Development, c1997 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://www.acdlabs.com/iupac/nomenclature/>>.

38. *Project : Macromolecular Division (IV) : Compendium of macromolecular terminology and nomenclature (The Purple Book), 2nd Edition* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2008, page last modified 12 December 2008 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://media.iupac.org/projects/2002/2002-048-1-400.html>>.
39. *Compendium of Analytical Nomenclature : Definitive Rules 1997* [online]. Web site constructed by David S. Moore. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, last updated 31 July 2002 [cit. 2009-09-03]. Dostupné na World Wide Web: <http://media.iupac.org/publications/analytical_compendium/>.
40. *Biochemical Nomenclature and Related Documents : "White Book"* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://www.chem.qmul.ac.uk/iupac/bibliog/white.html>>.
41. *KTD - Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online databáze]. Praha : Knihovnický institut NK ČR, 2003- [cit. 2010-02-23]. Dostupná na World Wide Web: <<http://sigma.nkp.cz/cze/kt/>>.
42. format. In DAINTITH, John; WRIGHT, Edmund (ed.). 2008. *A Dictionary of Computing* [online]. Oxford (Velká Británie) : Oxford University Press, 2008 [cit. 2010-06-12]. Pro oprávněné uživatele dostupné v digitální knihovně Oxford Reference Online: <<http://www.oxfordreference.com/>>.
43. BERNERS-LEE, Tim. *Metadata Architecture* [online]. W3C, January 1997, Last edit 2009/08/27 21:38:08 [cit. 2019-06-12]. Dostupné na World Wide Web: <<http://www.w3.org/DesignIssues/Metadata.html>>.
44. GRUBER, Thomas R. A translation approach to portable ontology specifications. *Knowledge Acquisition* [online]. 1993, vol. 5, issue 2 [cit. 2010-06-02], s. 199-220. Pro oprávněné uživatele dostupné v digitální knihovně ScienceDirect: <<http://www.sciencedirect.com/>>. Preprint dostupný na World Wide Web: <<http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>>.
45. recursion. In DAINTITH, John; WRIGHT, Edmund (ed.). *A Dictionary of Computing* [online]. Oxford (Velká Británie) : Oxford University Press, 2008 [cit. 2010-05-01]. Pro oprávněné uživatele dostupné v digitální knihovně Oxford Reference Online: <<http://www.oxfordreference.com/>>.
46. BERNERS-Lee, Tim. *Glossary* [online]. Cambridge (MA, USA) : Tim Berners-Lee, Last modified: Fri Jul 23 00:26:26 EDT 1999 [cit. 2010-06-12]. Semantic Web. Slovník ke knize Weaving the Web : The original design and ultimate destiny of the World Wide Web (ISBN 0-062-51587-X). Dostupné na World Wide Web: <<http://www.w3.org/People/Berners-Lee/Weaving/glossary.html>>.
47. GARFIELD, Eugene. *The Agony and the Ecstasy - The History and Meaning of the Journal Impact Factor* [online]. September 16, 2005 [cit. 2008-09-06]. 22 s. Připraveno pro International Congress on Peer Review and Biomedical Publication, Chicago, September 16, 2005. Dostupné na World Wide Web: <<http://garfield.library.upenn.edu/papers/jifchicago2005.pdf>>.
48. HIRSCH, J. E. An index to quantify an individual's scientific research output. *PNAS : Proceedings of the National Academy of Sciences of the United States of America* [online]. November 15, 2005, vol. 102, no. 46 [cit. 2010-05-15], s. 16569-16572. Dostupné na World Wide Web: <<http://www.pnas.org/content/102/46/16569.full.pdf>>. doi:10.1073/pnas.0507655102. ISSN 0027-8424 (tištěná verze). ISSN 1091-6490 (online verze).
49. MERTA, Augustin. *Společenské aspekty komunikace odborných informací*. Vyd. 1. Praha : ÚVTEI, 1970. 209 s. Dostupné také v elektronické podobě na CD-ROM Informační studia a

knihovnictví v elektronických textech I. (Praha : Univerzita Karlova, c2001) a na World Wide Web v rámci Elektronických studijních textů: <<http://texty.jinonice.cuni.cz/studijni-texty/archiv-starsich-textu/021.pdf>>.

50. KÖNIGOVÁ, Marie. Vybrané kapitoly z informační vědy. In PAPIK, Richard; SOUČEK, Martin; STÖCKLOVÁ, Anna (ed.). *Informační studia a knihovnictví v elektronických textech* [CD-ROM]. Praha : Univerzita Karlova, c2001. 18 s. Dostupné také na World Wide Web v rámci Elektronických studijních textů: <<http://texty.jinonice.cuni.cz/studijni-texty/archiv-starsich-textu/019.pdf>>.
51. BRATKOVÁ, Eva. *Metody citování literatury a strukturování bibliografických záznamů podle mezinárodních norem ISO 690 a ISO 690-2 : metodický materiál pro autory vysokoškolských kvalifikačních prací* [online]. Verze 2.0, aktualiz. a rozšíř. Praha : Odborná komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací, Asociace knihoven vysokých škol ČR, 2008-12-22 [2009-11-02]. 60 s. Dostupné na World Wide Web: <<http://www.evskp.cz/SD/4c.pdf>>.
52. ČSN ISO 690. *Dokumentace - Bibliografické citace - Obsah, forma a struktura*. Praha : Český normalizační institut, 1996. 32 s.
53. ČSN ISO 690-2. *Informace a dokumentace – Bibliografické citace – Část 2 : Elektronické dokumenty nebo jejich části*. Praha : Český normalizační institut, 2000. 22 s.
54. BRATKOVÁ, Eva. *Bibliografické odkazy pro seznamy a citace : Oficiální výukové stránky Ústavu informačních studií a knihovnictví FF UK* [online]. Praha : ÚISK FF UK, c1996-2009, Created: 1996-11-06, Last updated: 2009-03-11 [cit. 2009-11-02]. Dostupné na World Wide Web: <<http://www1.cuni.cz/~brt/bibref/bibref.html>>.
55. KUBÁLKOVÁ, Petra. *Instrukce k citování literatury dostupné v prostředí WWW na vysokých školách : analýza a hodnocení na základě standardu ČSN ISO 690 a ČSN ISO 690-2 [Guidelines for citing from Resources available on web pages of Universities in the Czech Republic : Analysis and Assessment Based on Standards CSN ISO 690 a CSN ISO 690-2]*. Praha, 2006-04-05. 39 s., VI s. příl. Bakalářská práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí bakalářské práce PhDr. Eva Bratková.
56. POLÁKOVÁ, Barbora. *Americké citační styly [American citation styles]*. Praha, 2007-07-11. 43 s. VIII s. příl. Bakalářská práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí bakalářské práce PhDr. Eva Bratková.
57. VALÍČEK, Jan. *Informační zdroje a informační zabezpečení oboru zoopaleontologie se zaměřením na třídu trilobitů a způsoby a formy citování literatury*. Praha, 2004. 180 s., 12 s. příl. Rigorózní práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí rigorózní práce PhDr. Eva Bratková.
58. TKAČÍKOVÁ, Daniela. *JAK zpracovávat bibliografické citace podle ČSN ISO 690-2* [online]. Ostrava : Ústřední knihovna Vysoké školy báňské-Technické univerzity Ostrava, c1998-2009, Aktualizováno 2008-03-24 [cit. 2009-11-02]. Dostupné na World Wide Web: <<http://knihovna.vsb.cz/kurzy/citace/ele01.html>>.
59. MANSFIELD, Jerry W. Why Not a Uniform Citation System? *The Journal of Academic Librarianship*. 1984, vol. 10, no. 4, s. 220-222. ISSN 0099-1333.
60. NIENHAUS, Brian. Helping Students Improve Citation Performance. *Business Communication Quarterly*. 2004, vol. 67, s. 337-348. Dostupné také na World Wide Web: <<http://www.sagepub.com/holt/articles/Nienhaus.pdf>>. doi:10.1177/1080569904268123. ISSN 1080-5699 (tištěná verze). ISSN 1552-4191 (online verze).
61. OPPENHEIM, Charles; SMITH, Richard. Student citation practices in an Information Science Department. *Education for Information*. 2001, vol. 19, no. 4, s. 299-323. Pro oprávněné

uživatelé dostupné také v digitální knihovně IOS Press: <<http://iospress.metapress.com/content/xkqp43pfnweb0fyn/>>. ISSN 0167-8329 (tištěná verze). ISSN 1875-8649 (online verze).

62. CLARKE, Maria Elizabeth; OPPENHEIM, Charles. Citation behaviour of information science students II : Postgraduate students. *Education for Information*. 2006, vol. 24, no. 1, s. 1-30. Pro oprávněné uživatele dostupné také v digitální knihovně IOS Press: <<http://iospress.metapress.com/content/98a5akqrrfeyvgpa/>>. ISSN 0167-8329 (tištěná verze). ISSN 1875-8649 (online verze).
63. BALÍKOVÁ, Marie. Harmonizace věcné katalogizace v Česku : sen či realita? In *Knihovny současnosti 2004 : sborník z 12. konference, konané ve dnech 14.-16. září 2004 v Seči u Chrudimi*. Brno : Sdružení knihoven ČR, 2004, s. 307-322. Dostupné také na World Wide Web: <<http://www.sdruk.cz/sec/2004/sbornik/2004-2-307.pdf>>. ISBN 80-86249-30-1.
64. HORÁK, Pavel; KOTRLOVÁ, Jindřiška; SKOLKOVÁ, Linda. Zpráva ze semináře CASLIN 2008. *Ikaros* [online]. 2008, roč. 12, č. 8 [cit. 2008-09-21]. Dostupné na World Wide Web: <<http://www.ikaros.cz/node/4937>>. URN-NBN:cz-ik4937. ISSN 1212-5075.
65. Ikaros, redakce. 2008. Hlubková indexace: inovace ve vyhledávání dokumentů pro odbornou veřejnost (Jim McGinty). *Ikaros* [online]. 2008, roč. 12, č. 5/2 [cit. 2008-09-21]. Dostupné na World Wide Web: <<http://www.ikaros.cz/node/4780>>. URN-NBN:cz-ik4780. ISSN 1212-5075.
66. GREEN, Toby. *We Need Publishing Standards for Datasets and Data Tables* [online]. Paris : OECD Publishing, 2009 [cit. 2009-05-20]. 29 s. OECD Publishing White Paper. Dostupné na WWW: <<http://dx.doi.org/10.1787/603233448430>>. doi: 10.1787/603233448430.
67. LAWRENCE, Steve; LEE GILES, C.; BOLLACKER, Kurt. Digital libraries and autonomous citation indexing. *Computer*. 1999, vol. 32, no. 6, s. 67-71. Preprint dostupný na World Wide Web: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.1607>>.
68. CORTEZ, Eli, et al. A Flexible Approach for Extracting Metadata From Bibliographic Citations. *Journal of the American Society for Information Science and Technology*. 2009, vol. 60, no. 6, s. 1144-1158. Pro oprávněné uživatele dostupné také v digitální knihovně Wiley InterScience: <<http://www3.interscience.wiley.com/>>.
69. RAYMOND, Eric Steven. *The Cathedral and the Bazaar* [online]. Version 3.0. Thyrsus Enterprises, c2000 [cit. 2008-04-02]. Dostupné na World Wide Web: <<http://www.catb.org/estr/writings/cathedral-bazaar/cathedral-bazaar/>>.
70. RAYMOND, Eric Steven. *Katedrála a tržiště* [online]. Překlad do češtiny Miloslav Nič. Zvon.org, 27.7.1999 [cit. 2008-04-02]. Dostupné na World Wide Web: <http://www.zvon.org/ZvonHTML/Translations/cathedral-bazaar/front_cs.html>.
71. American Psychological Association. *APA Style* [online]. Washington (DC, USA) : American Psychological Association, c2010 [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://apastyle.apa.org/>>.
72. *The Chicago Manual of Style Online* [online]. 15th ed. The University of Chicago, c2006, 2007 [cit. 2010-06-20]. Dostupné na World Wide Web: <<http://www.chicagomanualofstyle.org/home.html>>.
73. *AMA Manual of Style : a guide for authors and editors* [online]. American Medical Association, 2010 [cit. 2010-06-20]. Dostupné na World Wide Web: <<http://www.amamanualofstyle.com/oso/public/index.html>>.
74. *Journal Citation Reports* [online]. Thomson Reuters, c2010 [cit. 2010-05-01]. Pro oprávněné uživatele dostupné v rámci platformy Web of Knowledge: <<http://isiknowledge.com/>>.

75. BRATKOVÁ, Eva. Citace odborné literatury jako nástroj rozvoje služeb a integrace digitálních knihoven. In *AKP 2001 : Automatizace knihovnických procesů - 8 : sborník z 8. ročníku semináře pořádaného ve dnech 24.-25. dubna 2001 v Liberci*. Praha : ČVUT, Výpočetní a informační centrum, 2001, s. 109-120. Dostupné také na World Wide Web: <<http://knihovny.cvut.cz/akp/clanky/12.pdf>>.
76. BRATKOVÁ, Eva. Otevřený přístup, digitální knihovny a citační služby. In *Inforum 2006 – 12. konference o profesionálních informačních zdrojích, Praha 23. – 25.5.2006* [online]. Praha : AiP, 2006 [cit. 2009-11-02]. 21 s. Dostupné na World Wide Web: <http://www.inforum.cz/inforum2006/pdf/Bratkova_Eva.pdf>.
77. HARDY, Rachel, et al. *Open Access Citation Information. Final Report – Extended Version*. JISC Committee for the Information Environment (JCIE) Scholarly Communications Group, September 2005 [cit. 2009-11-02]. Dostupné na World Wide Web: <http://www.jisc.ac.uk/uploaded_documents/iv%20OA%20Citation%20Information%20FINAL%20Extended%20Report.DOC> nebo <<http://eprints.ecs.soton.ac.uk/11536/>> (v rámci ECS EPrints Repository).
78. HULL, Duncan; PETTIFER, Steve R.; KELL, Douglas B. Defrosting the Digital Library: Bibliographic Tools for the Next Generation Web. *PLoS Computational Biology* [online]. October 2008, vol. 4, issue 10 [cit. 2009-11-02], e1000204. Dostupné na World Wide Web: <<http://www.ploscompbiol.org/article/info:doi%2F10.1371%2Fjournal.pcbi.1000204>>. doi:10.1371/journal.pcbi.1000204.
79. Template:Citation. In *Wikipedia, the free encyclopedia* [online]. San Francisco (CA, USA) : Wikimedia Foundation, This page was last modified on 19 October 2009 at 23:29 [cit. 2009-11-02]. Dostupné na World Wide Web: <<http://en.wikipedia.org/wiki/Template:Citation>>.
80. NIELSEN, Finn Årup. Scientific citations in Wikipedia. *First Monday* [online]. August 2007, vol. 12, no. 8 [cit. 2009-11-02]. Dostupné na World Wide Web: <http://firstmonday.org/issues/issue12_8/nielsen/index.html>.
81. KOTRLOVÁ, Jindřiška; SKOLKOVÁ, Linda. Zpráva ze semináře CASLIN 2007. *Ikaros* [online]. 2007, roč. 11, č. 9 [cit. 2008-09-21]. Dostupné na World Wide Web: <<http://www.ikaros.cz/node/4305>>. URN-NBN:cz-ik4305. ISSN 1212-5075.
82. JEDLIČKOVÁ, Petra. *Česká republika v procesu transformace: Globalizace, informační politiky v ČR, EU a USA a odraz transformace v oblasti knihovnictví a informační vědy*. Praha, 2007. 186 s. Disertační práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí disertační práce doc. PhDr. Rudolf Vlasák.
83. BECK, Sheila; BECK, Richard. Web Citation : A Proposal for Standardized Specification. *Online* [online]. July-August 2006, vol. 30, no. 4 [cit. 2010-05-01], s. 31-34. Pro oprávněné uživatele dostupné na World Wide Web: <<http://search.ebscohost.com/>> nebo <<http://www.infotoday.com/online/jul06/index.shtml>>. ISSN 0146-5422.
84. ANDERSON, Benedict. *Imagined communities : reflections on the origin and spread of nationalism*. Revised ed. London; New York : Verso, c2006. xv, 240 s. ISBN 1-84467-086-4.
85. BILDER, Geoffrey. Citation Typing Ontology. *CrossTech* [online]. March 20, 2009 06:30 AM [cit. 2010-05-11]. Dostupné na World Wide Web: <http://www.crossref.org/CrossTech/2009/03/citation_typing_ontology.html>.
86. JOHNS, Adrian. *The Nature of the Book : Print and Knowledge in the Making*. Chicago; London : The University of Chicago Press, c1998. xxii, 753 s. ISBN 0-226-40122-7.
87. *Open Archives Initiative - Object Reuse and Exchange* [online]. OAI-ORE [cit. 2010-05-11]. Dostupné na World Wide Web: <http://www.openarchives.org/ore/ORE_Community.php>.

88. Ikaros, redakce. Přehled rámce pro výměnu a opětovné využití digitálních objektů v otevřených archivech OAI-ORE (Herbert Van de Sompel). *Ikaros* [online]. 2009, roč. 13, č. 5/2. [cit. 2009-11-02]. Dostupné na World Wide Web: <<http://www.ikaros.cz/node/5463>>. URN-NBN:cz-ik5463. ISSN 1212-5075.
89. marianner. Re: How do locations work? In *Kindle Boards* [online]. January 06, 2010, 06:44:49 PM [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://www.kindleboards.com/index.php?action=printpage;topic=17872.0>>.
90. FEBVRE, Lucien; MARTIN, Henri-Jean. *The Coming of the Book : The Impact of Printing 1450–1500*. Translated by David Gerard. London : Verso, 1990. ISBN 0-86091-797-5.
91. TRAVIS, Brian; HAHN, Michael. *HTML, SGML, PDF, XML: V čem se liší?* [online]. Překlad Miloslav Nič. 30.9.1999 [cit. 2010-05-11]. K dispozici též původní anglicky psaný text. Dostupné na World Wide Web: <http://old.zvon.org/ZvonHTML/Translations/HTMLSGMLPDFXMLdifference/front_all.html>.
92. SHOTTON, David. Semantic publishing: the coming revolution in scientific journal publishing. *Learned Publishing* [online]. April 2009, vol. 22, no. 2 [cit. 2010-06-09], s. 85-94. Dostupné na World Wide Web: <<http://dx.doi.org/10.1087/2009202>>.
93. CASSERLY, Mary F.; BIRD, James E. Web Citation Availability : A Follow-up Study. *Library Resources & Technical Services* [online]. 2008, vol. 52, no. 1 [cit. 2010-06-10], s. 42-53. Pro oprávněné uživatele dostupné na platformě EBSCOhost: <<http://www.ebscohost.com/>>. ISSN 0024-2527.
94. BERNERS-LEE, Tim, et al. A Framework for Web Science. *Foundations and Trends in Web Science*. 2006, vol. 1, issue 1, s. 1-130. Dostupné také na World Wide Web: <<http://eprints.ecs.soton.ac.uk/13347/1/1800000001%5B1%5D.pdf>>. doi:10.1561/1800000001. ISSN 1555-077X (tištěná verze). ISSN 1555-0788 (online verze).
95. Citation style language. In *Wikipedia, the free encyclopedia* [online]. San Francisco (CA, USA) : Wikimedia Foundation, This page was last modified on 23 March 2010 at 20:58 [cit. 2010-05-30]. Dostupné na World Wide Web: <http://en.wikipedia.org/wiki/Citation_style_language>.
96. *CSL : Citation Style Language* [online]. [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://xbiblio.sourceforge.net/csl/>>.
97. *CitationStyles.org* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://citationstyles.org/>>.
98. D'ARCUS, Bruce. *CSL Docs : Citation Style Language documentation* [online]. 2010-05-30 [cit. 2010-06-01]. Dostupné na World Wide Web: <<http://bitbucket.org/bdarcus/csl-docs>>.
99. *Zotero* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research, c2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.zotero.org/>>.
100. *MENDELEY : Research Networks* [online]. Beta 0.9. London : Mendeley [cit. 2010-06-09]. Dostupné na World Wide Web: <<http://www.mendeley.com/>>.
101. *Zotero Style Repository* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.zotero.org/styles>>.
102. CSL 0.8 Syntax Overview. In [*Zotero Documentation*] [online]. Washington (DC, USA) : George Mason University, Center for History and New Media, c2009, Last modified:

2010/03/23 16:51 [cit. 2010-06-09]. Dostupné na World Wide Web: <http://www.zotero.org/support/dev/csl_syntax_summary>.

103. *CitationStyles.org* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media, 2010 [cit. 2010-06-09]. CSL Schema. Dostupné na World Wide Web: <<http://citationstyles.org/citation-style-language/schema/>>.
104. ZELLE, Rintze M. *Citation Style Language 1.0 : Language Specification*. [online]. Contributors Frank G. Bennet, Jr., Bruce D'Arcus. Washington (DC, USA) : George Mason University, Center for History and New Media, 20100530 [cit. 2010-06-09]. Dostupné na World Wide Web: <<http://citationstyles.org/citation-style-language/schema/>>.
105. CHMELARĚ, Tomáš. *The numeric citation variant of the ISO 690(-2) style in CSL* [online]. Brno : Vysoké učení technické v Brně, Fakulta informačních technologií, 2010 [cit. 2010-04-30]. Dostupné na World Wide Web: <http://www.fit.vutbr.cz/research/view_product.php?id=123>.
106. *The Bibliographic Ontology* [online]. Co-editors Frédéric Giasson and Bruce D'Arcus. The Bibliographic Ontology, c2008-2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://bibliontology.com/>>.
107. D'ARCUS, Bruce; GIASSON, Frédéric. *Bibliographic Ontology Specification* [online]. Editor Frédéric Giasson. Revision: 1.3. Zitgist, c2006-2008, 4 November 2009 [cit. 2010-02-03]. Dostupné na World Wide Web: <<http://bibliontology.com/specification>>.
108. *ontology-browser : An OWL Ontology Browser* [online]. Manchester : The University of Manchester, c2007 [cit. 2010-05-15]. Dostupné na World Wide Web: <<http://code.google.com/p/ontology-browser/>>.
109. *The Bibliographic Ontology Website* [online]. Co-editors Frédéric Giasson and Bruce D'Arcus. The Bibliographic Ontology, c2008-2010 [cit. 2010-06-10]. The Bibliographic Ontology Usecase Examples. Dostupné na World Wide Web: <<http://bibliontology.com/examples>>.
110. *The Bibliographic Ontology Website* [online]. Co-editors Frédéric Giasson and Bruce D'Arcus. The Bibliographic Ontology, c2008-2010 [cit. 2010-06-10]. Projects. Dostupné na World Wide Web: <<http://bibliontology.com/projects>>.
111. *Creative Commons Attribution 1.0 Generic* [online]. Creative Commons [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://creativecommons.org/licenses/by/1.0/>>.
112. Mezinárodní federace knihovnických asociací a institucí. Cataloguing Section. *Functional Requirements for Bibliographic Records : Final Report — Current text* [online]. The Hague : IFLA, Latest revision: 26 December 2007 [cit. 2010-06-10]. Dostupné na World Wide Web: <http://archive.ifla.org/VII/s13/frbr/frbr_current_toc.htm>.
113. Dublin Core Metadata Initiative Usage Board. *DCMI Type Vocabulary* [online]. Singapore : DCMI, c1995-2010, Date Issued: 2008-01-14 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://dublincore.org/documents/dcmi-type-vocabulary/>>.
114. *SWAN Project : Ontology Ecosystem : SWAN Ontology* [online]. V. 1.2 (release candidate). SWAN Project, c2007-2009, last updated on Jan 17th, 2009 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://swan.mindinformatics.org/ontology.html>>.
115. SHOTTON, David. *CiTO, the Citation Typing Ontology, and its use for annotation of reference lists and visualization of citation networks* [online]. 2009 [cit. 2010-05-01]. 4 s. Připraveno pro 17th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB). Dostupné na World Wide Web: <http://imageweb.zoo.ox.ac.uk/pub/2008/publications/Shotton_ISMB_BioOntology_CiTO_final_postprint.pdf>.

116. McINNIS, Raymond G.; SYMES, Dal. David Riesman and the Concept of Bibliographic Citation. *College & Research Libraries*. September 1988, vol. 49, issue 5, s. 387-399.
117. GARFIELD, Eugene. Random Thoughts on Citationology. Its Theory and Practice. *Scientometrics* [online]. 1998, vol. 43, no. 1, s. 69-76 [cit. 2010-05-15]. Pro oprávněné uživatele dostupné v digitální knihovně SpringerLink: <<http://www.springerlink.com/>>. doi:10.1007/BF02458396. ISSN 0138-9130 (tištěná verze). ISSN 1588-2861 (online verze).
118. *Creative Commons Attribution 2.5 Generic* [online]. Creative Commons [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://creativecommons.org/licenses/by/2.5/>>.
119. SUAREZ-POTTS, Louis. *Interview: David Wilson, Bibliographic Project* [online]. OpenOffice.org, 2003-10-13 [cit. 2010-06-10]. Dostupné na World Wide Web: <http://www.openoffice.org/editorial/interview_dwilson.html>.
120. *bibliographic: Bibliographic Project Homepage* [online]. OpenOffice.org, Last updated 2008 November 4 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://bibliographic.openoffice.org/>>.
121. *bibliographic: Bibliographic Project Homepage* [online]. OpenOffice.org, Last updated 2008 November 4 [cit. 2010-06-10]. CiteProc. Dostupné na World Wide Web: <<http://bibliographic.openoffice.org/citeproc/index.html>>.
122. *CitationStyles.org* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media [cit. 2010-05-30]. CSL processors. Dostupné na World Wide Web: <<http://citationstyles.org/citation-style-language/csl-processors/>>.
123. BENNETT, Frank. *citeproc-js* [online]. This revision is from 2010-05-27 03:55 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://bitbucket.org/fbennett/citeproc-js/wiki/Highlights>>.
124. *Bibliographic/Developer Page* [online]. OpenOffice.org, November 4 2008 [cit. 2010-03-30]. Dostupné na World Wide Web: <http://wiki.services.openoffice.org/wiki/Bibliographic/Developer_Page>.
125. NIC, Miloslav. Chemical XML Formatting. In BANVILLE, Debra L. (ed.). *Chemical information mining : facilitating literature-based discovery*. Boca Raton (FL, USA) : Taylor & Francis, 2009, s. 99-119. ISBN 978-1-4200-7649-3.
126. *DocBook.org* [online]. [cit. 2010-05-01]. What is DocBook?. Dostupné na World Wide Web: <<http://docbook.org/whatis>>.
127. Organization for the Advancement of Structured Information Standards. *OASIS : Advancing open standards for the global information society* [online]. Billerica (MA, USA) : OASIS, c1993-2010 [cit. 2010-06-11]. About OASIS. Dostupné na World Wide Web: <<http://www.oasis-open.org/who/>>.
128. Organization for the Advancement of Structured Information Standards. *OASIS : Advancing open standards for the global information society* [online]. Billerica (MA, USA) : OASIS, c1993-2010 [cit. 2010-05-01]. OASIS DocBook TC. Dostupné na World Wide Web: <http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=docbook>.
129. *DocBook.org* [online]. [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.docbook.org/>>.
130. Organization for the Advancement of Structured Information Standards. *The DocBook Schema Version 5.0 : OASIS Standard* [online]. Billerica (MA, USA) : OASIS, 1 November 2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://docs.oasis-open.org/docbook/specs/docbook-5.0-spec.html>>.

131. *DocBook V5.0 RELAX NG* [online]. Billerica (MA, USA) : OASIS [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.docbook.org/xml/5.0/rng/>>.
132. WALSH, Norman. *DocBook 5: The Definitive Guide* [online]. Edited by Richard Hamilton. Version 1.1 for DocBook V5.0+. Norman Walsh, c2010, Updated: 20 May, 2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://www.docbook.org/tdg5/en/html/>>.
133. *RefDB Homepage* [online]. Markus Hoenicka, c2004-2006, 2008-02-15 17:09:21 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://refdb.sourceforge.net/>>.
134. KOSEK, Jiří. *DocBook : Stručný úvod do tvorby a zpracování dokumentů* [online]. Revize 0.9. Jiří Kosek, c2001-2007, aktualizace 7.5.2007 [cit. 2010-05-30]. 4.11. Seznamy literatury. Dostupné na World Wide Web: <<http://www.kosek.cz/xml/db/ukazky.html#d4e1396>>.
135. KOSEK, Jiří. *DocBook : Stručný úvod do tvorby a zpracování dokumentů* [online]. Revize 0.9. Jiří Kosek, c2001-2007, aktualizace 7.5.2007 [cit. 2010-05-30]. 3.5 Inline elementy. Dostupné na World Wide Web: <<http://www.kosek.cz/xml/db/struktury.html#d4e727>>.
136. biblioentry. In WALSH, Norman. *DocBook 5: The Definitive Guide* [online]. Edited by Richard Hamilton. Version 1.1 for DocBook V5.0+. Norman Walsh, c2010, Updated: 20 May, 2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://www.docbook.org/tdg5/en/html/biblioentry.html>>.
137. Organization for the Advancement of Structured Information Standards. *Index of /docbook/rng/ 5.0/* [online]. Billerica (MA, USA) : OASIS, Last Modified: 06-Feb-2008 08:13 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://docs.oasis-open.org/docbook/rng/5.0/>>.
138. *OpenDocument XML.org* [online]. Billerica (MA, USA) : OASIS, c1993-2010 [cit. 2010-05-30]. History of OpenDocument. Dostupné na World Wide Web: <<http://opendocument.xml.org/milestones>>.
139. Organization for the Advancement of Structured Information Standards. *Open Document Format for Office Applications (OpenDocument) Version 1.2* [online]. Part 1: Introduction and OpenDocument Schema. Committee Draft 04. Billerica (MA, USA) : OASIS, 15 December 2009 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://docs.oasis-open.org/office/v1.2/OpenDocument-v1.2-part1.html>>.
140. *OpenOffice.org 3.0 New Features* [online]. OpenOffice.org [cit. 2010-05-30]. Dostupné na World Wide Web: <http://www.openoffice.org/dev_docs/features/3.0/>.
141. International Digital Publishing Forum. *Open Publication Structure (OPS) 2.0 v1.0* [online]. Toronto (Kanada) : International Digital Publishing Forum, c2007, September 11, 2007 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://www.idpf.org/2007/ops/ops2.0/download/>>.
142. International Digital Publishing Forum. *Open Packaging Format (OPF) 2.0 v1.0* [online]. Toronto (Kanada) : International Digital Publishing Forum, c2007, September 11, 2007 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://www.idpf.org/2007/opf/opf2.0/download/>>.
143. International Digital Publishing Forum. *IDPF : International Digital Publishing Forum* [online]. Toronto (Kanada) : IDPF [2009-09-03]. Specifications. Dostupné na World Wide Web: <<http://idpf.org/specs.htm>>.
144. Library of Congress. Network Development and MARC Standards Office. *MARC Code List for Relators* [online]. Washington (DC, USA) : Library of Congress, May 13, 2010 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://www.loc.gov/marc/relators/>>.

145. Text Encoding Initiative. *TEI: Text Encoding Initiative* [online]. Charlottesville (VA, USA) : TEI Consortium [cit. 2010-06-10]. About the TEI. TEI: History. Dostupné na World Wide Web: <<http://www.tei-c.org/About/history.xml>>.
146. Text Encoding Initiative. *P5: Guidelines for Electronic Text Encoding and Interchange* [online]. Version 1.6.0. Charlottesville (VA, USA) : TEI Consortium, 2007, Last updated on February 12th 2010. This page generated on 2010-02-13T11:02:50Z [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index-toc.html>>.
147. BURNARD, Lou; RAHTZ, Sebastian. *Text Encoding Initiative* [online]. 2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://sourceforge.net/projects/tei/>>.
148. Text Encoding Initiative. *P5: Guidelines for Electronic Text Encoding and Interchange* [online]. Version 1.6.0. Charlottesville (VA, USA) : TEI Consortium, 2007, Last updated on February 12th 2010. This page generated on 2010-02-13T11:02:50Z [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index-toc.html>>.
149. MURRAY-RUST, Peter; RZEPA, Henry S. *Chemical Markup Language. A Position Paper.* [online]. 2001-04-10 [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://cml.sourceforge.net/historical/position.html>>.
150. JIRAT, Jiri. *Chemical Markup Language 1.0 reference with examples* [online]. [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://www.zvon.org/xxl/CML1.0/Output/index.html>>.
151. MURRAY-RUST, Peter. *The CML Blog* [online]. September 2006- [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://wwmm.ch.cam.ac.uk/blogs/cml/>>.
152. *cml.sourceforge.net - OpenSource Site for CML* [online]. Updated: May 15, 2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://cml.sourceforge.net/>>.
153. billyfish et al. *Chemical Markup Language* [online]. 2010 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://sourceforge.net/projects/cml/>>.
154. *XML-based IUPAC Standard for Experimental and Critically Evaluated Thermodynamic Property Data Storage and Capture* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2006, Page last modified 19 December 2006 [cit. 2010-06-11]. Číslo projektu: 2002-055-3-024. Dostupné na World Wide Web: <<http://media.iupac.org/projects/2002/2002-055-3-024.html>>.
155. *ThermoML version 3.0* [online]. [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://media.iupac.org/namespaces/ThermoML/ThermoML.xsd>> nebo <<http://trc.nist.gov/ThermoML.xsd>>.
156. International Union of Pure and Applied Chemistry. Committee on Printed and Electronic Publications. XML-based IUPAC standard for experimental, predicted, and critically evaluated thermodynamic property data storage and capture (ThermoML) (IUPAC Recommendations 2006). Prepared for publication by Michael Frenkel et al. *Pure and Applied Chemistry* [online]. 2006, vol. 78, no. 3 [cit. 2010-06-11], s. 541-612. Dostupné na World Wide Web: <<http://www.iupac.org/publications/pac/78/3/0541/>>. doi:10.1351/pac200678030541.
157. *ThermoML : An XML-Based IUPAC Standard for Storage and Exchange of Experimental Thermophysical and Thermochemical Property Data* [online]. Boulder (CO, USA) : National Institute of Standards and Technology, Thermophysical Properties Division, TRC Group [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://trc.nist.gov/ThermoML.html>>.
158. National Institute of Standards and Technology (Spojené státy americké). *TRC Group* [online]. Boulder (CO, USA) : National Institute of Standards and Technology, Thermophysical

Properties Division, TRC Group [cit. 2010-05-01]. TRC History. Dostupné na World Wide Web: <<http://trc.nist.gov/history.html>>.

159. *IUPAC Compendium of Chemical Terminology – the Gold Book* [online]. XML on-line corrected version created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c2005-2009 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://goldbook.iupac.org/>>. doi:10.1351/goldbook.
160. *Project: Standard XML data dictionaries for chemistry* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, Committee on Printed and Electronic Publications, 2002-2006 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://www.iupac.org/web/ins/2002-022-1-024>>.
161. VLASÁK, Rudolf; BULÍČKOVÁ, Soňa. *Základy projektování informačních systémů*. 1. vyd. Praha : Karolinum, 2003. 144 s. ISBN 80-246-0727-1.
162. CAPLAN, Priscilla. *Metadata Fundamentals for All Librarians*. Chicago : American Library Association, 2003. ix, 192 s. ISBN 0-8389-0847-0.
163. Unicode Consortium. *The Unicode Standard* [online]. Edited by Julie D. Allen. Version 5.2. Mountain View (CA, USA) : Unicode, December 2009 [cit. 2010-03-16]. Chapter 16. Dostupné na World Wide Web: <<http://www.unicode.org/versions/Unicode5.2.0/ch16.pdf>>.
164. Unicode Consortium. *Unicode 5.2 Character Code Charts* [online]. Mountain View (CA, USA) : Unicode, c1991-2010, Last updated: - 2. březen 2010 0:35:41 [cit. 2010-03-16] Controls: C0. Dostupné na World Wide Web: <<http://www.unicode.org/charts/PDF/U0000.pdf>>.
165. GOYVAERTS, Jan. *Regex Tutorial* [online]. c2003-2010, Page last updated: 17 June 2009 [cit. 2010-03-16]. Literal Characters and Special Characters. Dostupné na World Wide Web: <<http://www.regular-expressions.info/characters.html>>.
166. *RIS Format Specifications* [online]. Thomson Reuters, This page was last modified on: February 14, 2001 [cit. 2010-03-25]. Dostupné na World Wide Web: <http://www.refman.com/support/risformat_intro.asp>.
167. CHAMIER, Richard. *Case #TS-00136761:Export format specification:ref:00D3un.50079f5jA:ref* [elektronická pošta]. Message to: Linda Jansova. Po 10.8.2009 17:08 [cit. 2010-02-03]. Osobní komunikace.
168. *RefWorks* [online]. RefWorks, c2009 [cit. 2010-03-30]. About Us. Dostupné na World Wide Web: <http://www.refworks.com/content/about_us.asp>.
169. *EndNote* [online]. Thomson Reuters [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://endnote.com/>>.
170. *EndNote Web* [online]. Thomson Reuters [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://endnoteweb.com/>>.
171. *EndNote* [online]. Thomson Reuters [cit. 2010-03-02]. Technical Support & Services. Dostupné na World Wide Web: <<http://www.endnote.com/support/>>.
172. *[DTD XML verze formátu EndNote]* [online]. Thomson Reuters, EndNote [cit. 2010-03-15]. Dostupné na World Wide Web: <<http://www.endnote.com/support/helpdocs/RFXML.zip>>.
173. PATASHNIK, Oren. *BibTeXing* [online]. February 8, 1988 [cit. 2010-04-09]. 16 s. Dostupné na World Wide Web: <<http://www.ctan.org/get/biblio/bibtex/contrib/doc/btxdoc.pdf>>.
174. PATASHNIK, Oren. *Designing BibTeX styles* [online]. February 8, 1988 [cit. 2010-04-09]. 10 s. Dostupné na World Wide Web: <<http://www-users.cs.umn.edu/~thompson/latex/btxhak>>.

175. FEDER, Alexander. *BibTeX.org : Your BibTeX Resource* [online]. Vienna : Alexander Feder, c2006 [cit. 2010-04-09]. Dostupné na World Wide Web: <<http://www.bibtex.org/>>.
176. BibTeX. In *Wikipedia, the free encyclopedia* [online]. San Francisco (CA, USA) : Wikimedia Foundation, This page was last modified on 20 March 2010 at 04:16 [cit. 2010-04-09]. Dostupné na World Wide Web: <<http://en.wikipedia.org/wiki/BibTeX>>.
177. *LaTeX* [online]. This page was last modified on 8 April 2010, at 09:41 [cit. 2010-04-09]. Bibliography Management. Wikibooks, the open-content textbooks collection. Dostupné na World Wide Web: <http://en.wikibooks.org/wiki/LaTeX/Bibliography_Management>.
178. PATASHNIK, Oren. BibTeX yesterday, today, and tomorrow. *TUGboat* [online]. 2003, vol. 24, no. 1 - Proceedings of the 2003 Annual Meeting [cit. 2010-04-09], s. 25-30. Dostupné na World Wide Web: <<http://www.tug.org/TUGboat/Articles/tb24-1/patashnik.pdf>>.
179. Unicode Consortium. *The Unicode Consortium* [online]. Mountain View (CA, USA) : Unicode, c1991-2010, Last updated: - 13. únor 2009 0:32:59 [cit. 2010-06-15]. Chronology of Unicode Version 1.0. Dostupné na World Wide Web: <<http://www.unicode.org/history/versionone.html>>.
180. *LyX wiki* [online]. Page last modified on 2010-02-21 13:46 CET [cit. 2010-06-15]. BibTeX. Encoding. Dostupné na World Wide Web: <<http://wiki.lyx.org/BibTeX/Tips#encoding>>.
181. ‘String too long’ in BibTeX. In *The UK TeX Archive* [online]. last modified on 2010-04-07 [cit. 2010-06-15]. BibTeX. Encoding. Dostupné na World Wide Web: <<http://www.tex.ac.uk/cgi-bin/textfaq2html?label=bibstrtl>>.
182. *BibTeX style for the Czech references style* [online]. Original plainnat.bst by Patrick W Daly. Modifications by David Mudr{\a}k, last change Tue, 8 Apr 2008 19:41:58 +0000 [cit. 2010-06-15]. Dostupné na World Wide Web: <<http://repo.or.cz/w/csplainnat.git>>.
183. MARTÍNEK, David. *LaTeXové speciality: Český styl pro BibTeX* [online]. Poslední modifikace: 22. října 2008 [cit. 2010-06-15]. Dostupné na World Wide Web: <<http://www.fit.vutbr.cz/~martinek/latex/czechiso.html.cs.iso-8859-2>>.
184. LURATI, Brenno; PREVITALI, Luca. *BibTeXML* [online]. 2001 [cit. 2010-06-10]. 68 s. Institut für Technische Informatik und Kommunikationsnetze; Eidgenössische Technische Hochschule Zürich. Diplomová práce. Tutoři diplomové práce Dr. Erik Wilde, Dr. Bernhard Plattner. Vedoucí diplomové práce Dr. Bernhard Plattner. Dostupné na World Wide Web: <<http://dret.net/netdret/docs/da-ws2000-lurati-previtali.pdf>>.
185. *BibTeXML* [online]. [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://sourceforge.net/projects/bibtexml/>>.
186. KNOUF, Nick. *bibTeX Definition in Web Ontology Language (OWL) Version 0.1* [online]. Working Draft. MIT, c2004 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://zeitkunst.org/bibtex/0.1/>>.
187. HELLMAN, Eric. *OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML* [online]. stable version 1.0. New Jersey (NJ, USA) : OCLC New Jersey [cit. 2010-06-21]. Dostupné na World Wide Web: <<http://ocoins.info/>>.
188. ANSI/NISO Z39.88-2004. *The OpenURL Framework for Context-Sensitive Services* [online]. Bethesda (MD, USA) : NISO, ANSI Approval Date 15 Apr 2005 [cit. 2010-06-21]. xiv, 104 s. Dostupné na World Wide Web: <<http://www.niso.org/standards/z39-88-2004/>>. ISBN 978-1-880124-61-1. ISSN 1041-5653.
189. *Application of Charset Registration* [online]. IANA [cit. 2010-06-21]. Dostupné na World Wide Web: <<http://www.iana.org/assignments/charset-reg/Big5-HKSCS>>.

190. Internet Assigned Numbers Authority. *CHARACTER SETS* [online]. IANA, last updated 2007-05-14 [cit. 2010-06-21]. Dostupné na World Wide Web: <<http://www.iana.org/assignments/character-sets>>.
191. SIMONSEN, K. *Character Mnemonics & Character Sets* [online]. June 1992 [cit. 2010-06-21]. 103 s. RFC (Request for Comments) series, 1345. Dostupné na World Wide Web: <<http://www.ietf.org/rfc/rfc1345.txt>>. ISSN 2070-1721.
192. YERGEAU, F. *UTF-8, a transformation format of ISO 10646* [online]. January 1998 [cit. 2010-06-21]. 10 s. RFC (Request for Comments) series, 2279. Dostupné na World Wide Web: <<http://www.ietf.org/rfc/rfc2279.txt>>. ISSN 2070-1721.
193. HELLMAN, Eric. *OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML* [online]. stable version 1.0. New Jersey (NJ, USA) : OCLC New Jersey [cit. 2010-06-21]. Brief guide to Implementing OpenURL 1.0 Context Object for Journal Articles. Dostupné na World Wide Web: <<http://ocoins.info/cobg.html>>.
194. BERNERS-LEE, Tim. *Uniform Resource Identifier (URI): Generic Syntax* [online]. January 2005 [cit. 2010-06-21]. 61 s. RFC (Request for Comments) series, 3986. Dostupné na World Wide Web: <<http://www.ietf.org/rfc/rfc3986.txt>>. ISSN 2070-1721.
195. *Registry for the OpenURL Framework - ANSI/NISO Z39.88-2004* [online]. NISO; OCLC [cit. 2010-05-15]. Core:Metadata Formats. Dostupné na World Wide Web: <http://alcme.oclc.org/openurl/servlet/OAIHandler?verb=ListRecords&metadataPrefix=oai_dc&set=Core:Metadata+Formats>.
196. Library of Congress. *MARCXML : MARC 21 XML Schema : Official Web Site* [online]. Washington (DC, USA) : Library of Congress, March 13, 2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/marcxml/>>.
197. KEITH, Corey. *MARCXML: The MARC 21 XML Schema* [online]. Version 1.2. Washington (DC, USA) : Library of Congress, May 21, 2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>>.
198. Library of Congress. Network Development and MARC Standards Office. *Abbreviated Title Source Codes* [online]. Washington (DC, USA) : Library of Congress, April 6, 2010 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/sourcelist/abbreviated-title.html>>.
199. Library of Congress. *MODS : Metadata Object Description Schema Official Website* [online]. Washington (DC, USA) : Library of Congress, April 6, 2010 [cit. 2010-04-10]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/mods/>>.
200. Library of Congress. *[Záznam časopiseckého článku ve formátu MODS verze 3.0]* [online]. Washington (DC, USA) : Library of Congress [cit. 2010-05-15]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/mods/v3/modsjournal.xml>>.
201. DLF Aquifer. *Digital Library Federation / Aquifer Implementation Guidelines for Shareable MODS Records* [online]. Version 1.1. Digital Library Federation, March 2009 [cit. 2010-05-15]. 117 s. Dostupné na World Wide Web: <https://wiki.dlib.indiana.edu/confluence/download/attachments/24288/DLFMODS_ImplementationGuidelines.pdf>.
202. TitlePractices. In *Best Practices for OAI Data Provider Implementations and Shareable Metadata* [online]. Digital Library Federation; National Science Digital Library, This page was last modified 00:28, 27 June 2007 [cit. 2010-05-15]. Dostupné na World Wide Web: <<http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/TitlePractices>>.

203. Library of Congress. *MODS : Metadata Object Description Schema Official Website* [online]. Washington (DC, USA) : Library of Congress, April 6, 2010 [cit. 2010-04-10]. Guidance. Dostupné na World Wide Web: <<http://www.loc.gov/standards/mods/mods-guidance.html>>.
204. Library of Congress. Network Development and MARC Standards Office. *Network Source Codes for Vocabularies, Rules, and Schemes* [online]. Washington (DC, USA) : Library of Congress, Network Development and MARC Standards Office, June 10, 2009 [cit. 2010-06-11]. Name and title Authority Source Codes. Dostupné na World Wide Web: <<http://www.loc.gov/standards/sourcelist/name-title.html>>.
205. Library of Congress. *Extended Date/Time Format (EDTF)* [online]. Washington (DC, USA) : Library of Congress, March 29, 2010 [cit. 2010-05-15]. Dostupné na World Wide Web: <<http://www.loc.gov/standards/datetime/>>.
206. National Library of Medicine (Spojené státy americké). *Document Type Definitions (DTDs) for NLM Databases* [online]. Bethesda (MD, USA) : National Library of Medicine, Last reviewed: 19 January 2010, Last updated: 19 January 2010, First published: 17 October 2006 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.nlm.nih.gov/databases/dtd/>>.
207. National Library of Medicine (Spojené státy americké). *MEDLINE®/PubMed® XML Data Elements* [online]. Bethesda (MD, USA) : National Library of Medicine, Last reviewed: 15 April 2010, Last updated: 15 April 2010, First published: 24 February 2003 [cit. 2010-05-01]. Dostupné na World Wide Web: <http://www.nlm.nih.gov/bsd/licensee/data_elements_doc.html>.
208. National Library of Medicine (Spojené státy americké). *Journal Publishing Tag Set* [online]. Bethesda (MD, USA) : National Library of Medicine, Last updated: June 10, 2010 [cit. 2010-06-12]. Dostupné na World Wide Web: <<http://dtd.nlm.nih.gov/publishing/>>.
209. *Connotea: free online reference management for all researchers, clinicians and scientists* [online]. Nature Publishing Group, c2005-2010 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://connotea.org/>>.
210. *Zotero* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research, c2009 [cit. 2010-05-01]. About. Dostupné na World Wide Web: <<http://www.zotero.org/about/>>.
211. zotero data. In [*Zotero Documentation*] [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research, c2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <http://www.zotero.org/support/zotero_data>.
212. *Zotero* [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research, c2009 [cit. 2010-05-01]. Compatible sites. Dostupné na World Wide Web: <<http://www.zotero.org/translators/>>.
213. getting stuff into your library. In [*Zotero Documentation*] [online]. Washington (DC, USA) : George Mason University, Center for History and New Media Research, c2009 [cit. 2010-05-01]. Dostupné na World Wide Web: <http://www.zotero.org/support/getting_stuff_into_your_library>.
214. LUND, Ben, et al. Social Bookmarking Tools (II) : A Case Study - Connotea. *D-Lib Magazine* [online]. April 2005, vol. 11, no. 4 [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://www.dlib.org/dlib/april05/lund/04lund.html>>. doi:10.1045/april2005-lund. ISSN 1082-9873.
215. *Connotea: free online reference management for all researchers, clinicians and scientists* [online]. Nature Publishing Group, c2005-2010 [cit. 2010-05-01]. guide. Dostupné na World Wide Web: <<http://www.connotea.org/guide>>.

216. *Citace 2.0 - vše o citování literatury a dokumentů* (<http://www.citace.com/>) [online]. Citace.com, c2004-2010 [cit. 2010-05-01]. Dostupné na World Wide Web: <<http://www.citace.com/>>.
217. KRČÁL, Martin. Citace 2.0 – nové funkce, nové možnosti. In *Knihovny současnosti 2009 : sborník ze 17. konference, konané ve dnech 23.-25. června 2009 v Seči u Chrudimi*. Brno : Sdružení knihoven ČR, s. 99-103. Dostupné také na World Wide Web: <<http://www.sdruk.cz/sec/2009/sbornik/2009-3-099.pdf>>. ISBN 978-80-86249-54-4.
218. CiteULike. In *Wikipedia, the free encyclopedia* [online]. San Francisco (CA, USA) : Wikimedia Foundation, This page was last modified on 20 March 2010 at 17:02 [cit. 2010-04-09]. Dostupné na World Wide Web: <<http://en.wikipedia.org/wiki/CiteULike>>.
219. *CiteULike* [online]. Oversity [cit. 2010-05-01]. CiteULike Bookmarklets. Dostupné na World Wide Web: <<http://www.citeulike.org/bookmarklets.adp>>.
220. *CiteULike Help and FAQs* [online]. Oversity, This page was last modified on 13 January 2010, at 16:24 [cit. 2010-05-01]. Importing and Exporting. Dostupné na World Wide Web: <http://wiki.citeulike.org/index.php/Importing_and_Exporting>.
221. Národní knihovna České republiky. *Národní autority ČR* [online]. Praha : Národní knihovna ČR, c2007-2010 [cit. 2010-06-15]. Soubor základních formálních deskriptorů. Dostupné na World Wide Web: <<http://authority.nkp.cz/vecne-autority/soubor-formalnich-deskriptoru-1/soubor-zakladnich-formalnich-deskriptoru>>.
222. BRATKOVÁ, Eva. *Digitální knihovny s volným přístupem v oblasti vědy a výzkumu a identifikace a metadatový popis jejich objektů [Digital libraries with open access in science and research area and identification and metadata description of their objects]*. Praha, 2009-03-30. 242 s., 62 s. příl. Disertační práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Školitel doc. PhDr. Rudolf Vlasák.
223. ČSN ISO 5127. *Informace a dokumentace - Slovník*. Praha : Český normalizační institut, 2003. 160 s.
224. NOVÁKOVÁ, Marta. *Informatická typológia dokumentov : východiská, problémy, postupy*. Vyd. 1. Martin : Matica slovenská, 1981. 242 s.
225. BRATKOVÁ, Eva. K otázkám pojmu, třídění a typologie internetových a webovských informačních zdrojů. *Národní knihovna*. 1998, roč. 9, č. 5, s. 262-275. Aktualizovaný a doplněný postprint z dubna 2005 k dispozici v archivu E-LIS: <<http://eprints.rclis.org/3558/>>.
226. SCHWARZ, Josef. Nedokumentované informační prameny nebo informační objekty?. *Ikaros* [online]. 1999, roč. 3, č. 9 [cit. 1999-10-01]. Dostupný na World Wide Web: <<http://www.ikaros.cz/node/402>>. URN-NBN:cz-ik402. ISSN 1212-5075.
227. CEJPEK, Jiří. *Informace, komunikace a myšlení : úvod do informační vědy*. Praha : Karolinum, 2005. 233 s. ISBN 80-246-1037-X.
228. SCHAMBER, Linda. What Is a Document? Rethinking the Concept in Uneasy Times. *Journal of the American Society for Information Science*. 1996, vol. 47, no. 9, s. 669-671.
229. HILLESUND, Terje. Digital reading spaces: How expert readers handle books, the Web and electronic paper. *First Monday* [online]. 5 April 2010, vol. 15, no. 4 [cit. 2010-06-17]. Dostupné na World Wide Web: <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2762/2504>>.
230. SEARLE, John R. *Mysl, mozek a věda*. Vyd. 1. Praha : Mladá fronta, 1994. 129 s. Edice Váhy, sv. 13. ISBN 80-204-0509-7.
231. OTLET, Paul. *Traité de documentation : le livre sur le livre : théorie et pratique*. Bruxelles : Mondaneum, 1934. 431, [xii], viii s.

232. BUCKLAND, Michael K. Information as Thing. *Journal of the American Society for Information Science*. 1991, vol. 40, no. 5, s. 351-360.
233. LUND, Niels Windfeld; SKARE, Roswitha. Document Theory. In *Encyclopedia of Library and Information Sciences* [online]. 3rd ed. Taylor & Francis, 09 December 2009 [cit. 2010-06-15], s. 1632-1239. Pro oprávněné uživatele dostupné na World Wide Web: <<http://www.informaworld.com/>>. ISBN 978-0-8493-9711-0 (elektronická verze).
234. VICKERY, Brian Campbell; MILLS, Jack. *Fazetová klasifikácia. Návrh na triedenie knižničnej vedy*. Zostavil a preložil dr. Mikuláš Mičátek. Martin : Matica slovenská, 1968. 145 s.
235. JUSTICE, Alexander. Information Science as a Facet of the History of British Science: The Origins of the Classification Research Group. In RAYWARD, W. Boyd; BOWDEN, Mary Ellen (ed.). *The History and Heritage of Scientific and Technological Information Systems: Proceedings of the 2002 Conference: PDFs of proceedings* [online]. Medford (NJ, USA) : Information Today, 2004 [cit. 2010-06-15], s. 267-280. Vydáno pro American Society for Information Science and Technology a Chemical Heritage Foundation. ASIS&T Monograph Series. Dostupné na World Wide Web: <<http://www.chemheritage.org/pubs/asist2002/22-justice.pdf>>.
236. FREYTAG, Jurgen. On Typology of Information Sources and Means. *International Forum on Information and Documentation*. January 1978, vol. 3, no. 1, s. 3-7.
237. RÓZSA, György. Social sciences information: typology of sources. *UNESCO Bulletin for Libraries*. May-June 1978, vol. 32, no. 3, s. 167-171.
238. VICKERY, Brian C.; VICKERY, Alina. *Information science in theory and practice*. 3rd revised and enlarged ed. München : K. B. Saur, c2004. xiii, 400 s. ISBN 3-598-11658-6.
239. KAYE, David. Sources of information, formal and informal. *Library Management*. 1995, vol. 16, issue 5, s. 16-19. Pro oprávněné uživatele dostupné také na platformě ProQuest: <<http://proquest.umi.com/>>.
240. ROSSO, Mark A. User-Based Identification of Web Genres. *Journal of the American Society for Information Science and Technology* [online]. 2008, vol. 59, no. 7 [cit. 2010-05-30], s. 1053-1072. Pro oprávněné uživatele dostupné v digitální knihovně Wiley InterScience: <<http://www3.interscience.wiley.com/>>.
241. NOVÁKOVÁ, Marta. *Kapitoly z typovej klasifikácie dokumentov*. Bratislava : Slovenská technická knižnica, 1972. 134 s. Metodické pomôcky. D – študijné príručky, č. 19.
242. NOVÁKOVÁ, Marta. *Kapitoly z typovej klasifikácie dokumentov*. 2. vyd. Bratislava : Slovenská technická knižnica, 1976. 132 s. Metodické pomôcky. D – študijné príručky, č. 29.
243. VODIČKOVÁ, Hana. K pojmu „informační prameny“. *Bibliografický časopis*. 1968, roč. 1, č. 4, s. 133-140.
244. VÍTKOVÁ, Hana. *Informační prameny : jejich druhy a charakteristika*. Vyd. 1. Praha : ÚVTEI, 1973. 67 s. Učební texty.
245. BASLOVÁ-RANDOVÁ, Marie. *Vztah mezi uživatelem, typem informace a informačním pramenem*. Praha, 1983. 123, [47] s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Katedra vědeckých informací a knihovnictví. Vedoucí práce PhDr. Marie Königová, CSc.
246. SKOLEK, Jaroslav. Kniha a jiné dokumenty jako přímé zdroje informací. In CEJPEK, Jiří (sest.). *Využívání odborných informací*. Praha : ÚVTEI, Institut pro mimoškolní vzdělávání, 1983, s. 18-25. Metodická příručka pro lektory.

247. SEDLÁČKOVÁ, Beáta. *Dokumentografie. Část 1. Typologie dokumentů*. Opava : Slezská univerzita, 1993. 84 s.
248. SEDLÁČKOVÁ, Beáta; MARVANOVÁ. *Dokumentová komunikace*. 1. vyd. Praha : Národní knihovna ČR, 2007. 151 s. Studijní texty. ISBN 978-80-7050-535-9.
249. ŠILHÁNEK, Jaroslav. *Chemická informatika*. Praha : Vydavatelství VŠCHT, 2002. 355 s. Dostupné také na World Wide Web: <http://vydavatelstvi.vscht.cz/knihy/uid_isbn-80-7080-465-3/pages-img/>. ISBN 80-7080-465-3.
250. 250: , Registry for the OpenURL Framework - ANSI/NISO Z39.88-2004 , [online]. NISO; OCLC [cit. 2010-05-15]. Matrix defining the KEV Format to represent a journal publication. Dostupné na World Wide Web: <<http://alcm.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:journal>>.
251. *COinS Generator* [online]. version 1.7.3b. Openly Informatics, c1999-2005, May 21, 2008 [cit. 2010-06-21]. Dostupné na World Wide Web: <<http://generator.ocoins.info/>>.
252. Library of Congress. *MODS : Metadata Object Description Schema Official Website* [online]. Washington (DC, USA) : Library of Congress, April 9, 2010 [cit. 2010-04-10]. Outline of Elements and Attributes in MODS Version 3.3. Dostupné na World Wide Web: <<http://www.loc.gov/standards/mods/mods-outline.html>>.
253. National Library of Medicine (Spojené státy americké). *Medical Subject Headings* [online]. Bethesda (MD, USA) : National Library of Medicine, Last reviewed: 06 May 2010, Last updated: 06 May 2010, First published: 01 September 1999 [cit. 2010-06-11]. Dostupné na World Wide Web: <<http://www.nlm.nih.gov/mesh/>>.
254. National Library of Medicine (Spojené státy americké). *Medical Subject Headings* [online]. Bethesda (MD, USA) : National Library of Medicine, Last reviewed: 01 September 2009, Last updated: 01 September 2009, First published: 01 September 2009 [cit. 2010-06-11]. Publication Characteristics (Publication Types) - Scope Notes. Dostupné na World Wide Web: <<http://www.nlm.nih.gov/mesh/pubtypes.html>>.
255. Publication Types. In *PubMed Help* [online]. PubMed, 2005- , Last Update: June 7, 2010. [cit. 2010-06-08]. Dostupné na World Wide Web: <<http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helppubmed&part=pubmedhelp&rendertype=table&id=pubmedhelp.T42>>.
256. *Anglo-americká katalogizační pravidla : druhé vydání, revize 1988*. 1. české vyd. Praha : Národní knihovna, 1994. ISBN 80-7050-187-1.
257. Národní knihovna České republiky. *AACR2R/UNIMARC : Schválené české interpretace (duben 1999)* [online]. Praha : Národní knihovna ČR, duben 1999, 07.03.06 [cit. 2010-03-04]. Dostupné na World Wide Web: <http://www.nkp.cz/pages/page.php3?page=fond_def_int.htm>.
258. *Anglo-American Cataloging Rules : North American Text*. Prepared by The American Library Association, The Library of Congress, The Library Association and The Canadian Library Association. Chicago : American Library Association, 1967. xxi, 400 s.
259. Library of Congress. Cataloging Policy and Support Office. *Library of Congress Implementation of Amendments 2001 to AACR2* [online]. Washington (DC, USA) : Library of Congress, 11/28/2001 [cit. 2010-03-04]. Dostupné na World Wide Web: <<http://www.loc.gov/catdir/cpso/amen2001.html>>.
260. TILLET, Barbara B. AACR2's Updates for Electronic Resources : Response of a Multinational Cataloguing Code : A Case Study. In *Record of Workshop on Authority Control among Chinese, Korean and Japanese Languages held at National Institute of Informatics (NII) in cooperation with National Diet Library* [online]. Tokyo : National Institute of

- Informatics, March 2002 [cit. 2010-03-04], s. 165-173. Dostupné na World Wide Web: <<http://www.nii.ac.jp/publications/CJK-WS/cjk3-11a.pdf>>.
261. Library of Congress. *Cataloger's Desktop* [online]. Washington (DC, USA) : Library of Congress, c2009 [cit. 2010-03-04]. Pro oprávněné uživatele dostupné na World Wide Web: <<http://desktop.loc.gov/>>.
262. Joint Steering Committee for Development of RDA. *Joint Steering Committee for Development of RDA* [online]. Joint Steering Committee for Development of RDA, Last updated: 15 July 2009 [cit. 2010-03-04]. Historic Documents. Dostupné na World Wide Web: <<http://www.rda-jsc.org/docs.html>>.
263. *RDA : Full Draft* [online]. American Library Association, Canadian Library Association, The Chartered Institute of Library and Information Professionals, c2008 [cit. 2010-06-15]. Dostupné na World Wide Web: <http://www.rdatoolkit.org/constituencyreviewfiles/constituency_review.zip>.
264. *RDA: Resource Description & Access Toolkit* [online]. American Library Association; Canadian Library Association; CILIP c2010- [cit. 2010-06-15]. Dostupné na World Wide Web: <<http://www.rdatoolkit.org/>>.
265. DUNSIRE, Gordon. Distinguishing Content from Carrier : The RDA/ONIX Framework for Resource Categorization. *D-Lib Magazine* [online]. January/February 2007, vol. 13, no. 1/2 [cit. 2010-06-18]. Dostupné na World Wide Web: <<http://www.dlib.org/dlib/january07/dunsire/01dunsire.html>>. ISSN 1082-9873. doi:10.1045/january2007-dunsire.
266. *ONIX* [online]. London : EDItEUR, c2009 [cit. 2010-06-18]. Dostupné na World Wide Web: <<http://www.editeur.org/8/ONIX/>>.
267. *Mezinárodní desetinné třídění : zkrácené vydání*. Díl 1, Pomocné tabulky, Třídy 0-53. Praha : ARTIFEX, 1994. 177 s. Publikace UDC-P004 autorizovaná Konsorciem MDT licence UDC9306. ISBN 80-900034-6-X.
268. KOVÁŘ, Bohuslav. *Věcné pořádkání informací a selekční jazyky*. Vyd. 1. Díl 1. Úvod do problematiky, systematické pořádkání. Praha : ÚVTEI, 1981. 144 s. Učební texty, sv. 14.
269. BUCKLAND, Michael. *1895-2002 FID Achievements* [online]. Berkley (CA, USA) : University of California, School of Information [cit. 2010-03-06]. Dostupné na World Wide Web: <<http://people.ischool.berkeley.edu/~buckland/fidhist.html>>.
270. RAYWARD, W. Boyd. The International Federation for Information and Documentation (FID). In WIEGAND, Wayne A.; DAVIS, Don G., Jr. (ed.). *Encyclopedia of Library History*. New York : Garland Publ., 1994. 707 s. Heslo dostupné také na World Wide Web: <<http://people.lis.illinois.edu/~wrayward/otlet/FIDHIST2.htm>>. ISBN 0-8240-5787-2.
271. Mezinárodní federace knihovnických asociací a institucí. *Open Letter on IFLA and FID* [online]. The Hague : International Federation of Library Associations and Institutions, c1995-2001, Latest Revision: June 14, 2001 [cit. 2010-03-06]. V datu poslední revize opraven překlep (místo výrazu Juney je správné znění June). Dostupné na World Wide Web: <<http://archive.ifla.org/III/misc/fid-ifla.htm>>.
272. UDC Consortium. *UDC Consortium* [online]. The Hague : UDC Consortium, Last updated: 21 October 2009 [cit. 2010-03-06]. About Universal Decimal Classification. V datu poslední aktualizace opraven pravděpodobný překlep (původní rok 2010 opraven na rok 2009). Dostupné na World Wide Web: <<http://udcc.org/about.htm>>.
273. *Extensions and Corrections to the UDC* [online]. The Hague : UDC Consortium, Last updated: 29 January 2009 [2010-05-01]. 1x ročně. Dostupné na World Wide Web: <<http://udcc.org/ec.htm>>. ISSN 0014-5424.

274. UDC Consortium. *UDC Consortium* [online]. The Hague : UDC Consortium, Last updated: 21 February 2010 [cit. 2010-03-06]. Master Reference File (MRF). Dostupné na World Wide Web: <<http://udcc.org/mrf.htm>>.
275. UDC Consortium. *UDC Consortium* [online]. The Hague : UDC Consortium, Last updated: 01 September 2009 [cit. 2010-03-06]. Cancellations from the UDC Master Reference File. Dostupné na World Wide Web: <<http://udcc.org/cancellations.htm>>.
276. *Creative Commons Attribution-Share Alike 3.0 Unported* [online]. Creative Commons [cit. 2010-06-10]. Dostupné na World Wide Web: <<http://creativecommons.org/licenses/by-sa/3.0/>>.
277. Národní knihovna České republiky. *Mezinárodní desetinné třídění* [online]. 5. vyd. Praha : Národní knihovna ČR, říjen 2003 [cit. 2010-03-06]. Dostupné na World Wide Web: <<http://aip.nkp.cz/mdt/>>.
278. Mezinárodní unie pro čistou a užitou chemii. *IUPAC* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, last update: 02 September 2009 [cit. 2009-09-03]. Committee on Printed and Electronic Publications (CPEP). Dostupné na World Wide Web: <<http://iupac.org/web/ins/024>>.
279. NIČ, Miloslav. Osobní komunikace. 28. 8. 2009.
280. Mezinárodní unie pro čistou a užitou chemii. *IUPAC* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2007, page last modified [cit. 2009-08-17]. About IUPAC. Dostupné na World Wide Web: <<http://media.iupac.org/general/about.html>>.
281. Mezinárodní unie pro čistou a užitou chemii. *IUPAC : International Union of Pure and Applied Chemistry* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, last update: 02 September 2009 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://www.iupac.org/>>.
282. Mezinárodní unie pro čistou a užitou chemii. *IUPAC* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2008, page last modified 14 November 2008 [cit. 2009-09-03]. IUPAC Nomenclature Books Series. Dostupné na World Wide Web: <<http://media.iupac.org/publications/books/seriestitles/nomenclature.html>>.
283. Mezinárodní unie pro čistou a užitou chemii. *IUPAC* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2009, Page last modified 10 July 2009 [cit. 2009-09-01]. IUPAC Publications. e-resources. Dostupné na World Wide Web: <<http://media.iupac.org/publications/epub/index.html>>.
284. *Pure and Applied Chemistry* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, 1960- [cit. 2009-08-17]. Měsíčník. Dostupné na World Wide Web: <<http://www.iupac.org/publications/pac/index.html>>. ISSN 1365-3075 (online verze). ISSN 0033-4545 (tištěná verze).
285. *Pure and Applied Chemistry* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c2009 [cit. 2009-08-17]. Instructions for authors. Dostupné na World Wide Web: <http://www.iupac.org/publications/pac/instructions_for_authors.html>.
286. *Chemistry International* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, 1997- [cit. 2009-08-17]. Dvuměsíčník. Dostupné na World Wide Web: <<http://www.iupac.org/publications/ci/index.html>>. ISSN 0193-6484.
287. *Chemical Education International* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, Committee on Chemistry Education, 2000- [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://media.iupac.org/publications/cei/>>.

288. KOSATA, Beda. *Goldify* [online]. c2009 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://sourceforge.net/projects/goldify/>>.
289. *IUPAC Compendium of Chemical Terminology* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2006, page last modified 19 October 2006 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://media.iupac.org/publications/compendium/index.html>>.
290. *IUPAC Compendium of Chemical Terminology – the Gold Book* [online]. XML on-line corrected version created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c2005-2009 [cit. 2009-09-03]. Source documents. Dostupné na World Wide Web: <<http://goldbook.iupac.org/literature.html>>. doi:10.1351/goldbook.
291. *IUPAC Compendium of Chemical Terminology – the Gold Book* [online]. XML on-line corrected version created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c2005-2009 [cit. 2009-09-03]. About the IUPAC Compendium of Chemical Terminology (Gold Book). Dostupné na World Wide Web: <<http://goldbook.iupac.org/about.html>>. doi:10.1351/goldbook.
292. *Quantities, Units and Symbols in Physical Chemistry : - the Green Book -* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2007, Page last modified 4 September 2007 [cit. 2009-09-01]. Dostupné na World wide Web: <<http://media.iupac.org/publications/books/gbook/index.html>>.
293. *Nomenclature of Inorganic Chemistry : IUPAC Recommendations 2005* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2008, Page last modified 23 September 2008 [cit. 2009-09-01]. Dostupné na World wide Web: <<http://media.iupac.org/publications/books/author/connelly.html>>.
294. *Compendium of Macromolecular Nomenclature : The Purple Book, First Edition* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2006, page last modified 13 January 2006 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://media.iupac.org/publications/books/author/metanomski.html>>.
295. *Project : Macromolecular Division (IV) : Compendium of macromolecular terminology and nomenclature (The Purple Book), 2nd Edition* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2008, page last modified 12 December 2008 [cit. 2009-09-03]. Nomenclature of Regular Single-Strand Organic Polymers. Dostupné na World Wide Web: <<http://www.chem.qmul.ac.uk/iupac/rssop/>>.
296. *Compendium of Terminology and Nomenclature of Properties in Clinical Laboratory Sciences - The Silver Book* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, c1997-2001, page last modified 21 June 2001 [cit. 2009-09-03]. Dostupné na World Wide Web: <<http://media.iupac.org/publications/books/author/rigg.html>>.
297. JENKINS, Keith. Deciphering Journal Abbreviations with JAbbr. *Code4Lib* [online]. 2009-06-26, issue 7 [cit. 2010-05-30]. Dostupné na World Wide Web: <<http://journal.code4lib.org/articles/1758>>. ISSN 1940-5758.
298. *CAS Source Index (CASSI) Search Tool* [online]. Columbus (OH, USA) : American Chemical Society, Chemical Abstracts Service, c2009 [cit. 2010-06-20]. Dostupné na World Wide Web: <<http://cassi.cas.org/search.jsp>>.
299. *Numeric representation of dates and times (ISO 8601)* [online]. Prepared by: Miloslav Nic (Mila). [2010] [cit. 2010-06-23]. Dostupné na World Wide Web: <http://new.zvon.org/comp/tr/ref-Dates_and_Times.html>.

300. *Pure and Applied Chemistry* [online]. Research Triangle Park (NC, USA) : International Union of Pure and Applied Chemistry, 1960- [cit. 2009-08-17]. Měsíčník. Dostupné na World Wide Web: <<http://www.iupac.org/publications/pac/index.html>>. ISSN 0033-4545 (tištěná verze). ISSN 1365-3075 (online verze).

Přílohy

Příloha č. 1: Web of Science

1. Část záznamu časopiseckého článku ve formátu ISI (pro uložení v citačních manažerech EndNote, RefMan a ProCite):

```
FN ISI Export Format
VR 1.0
PT J
AU Welte, FJ
   Kim, SC
   Doshi, DJ
   O'Connor, SC
   Coughlin, BF
AF Welte, Frank J.
   Kim, Sunah C.
   Doshi, Devang J.
   O'Connor, Stephen C.
   Coughlin, Bret F.
TI Incorporation of a Formalized Emergency Radiology Curriculum to
   Facilitate Population of a MIRC-based Digital Teaching File
SO JOURNAL OF DIGITAL IMAGING
LA English
DT Article
DE Medical Imaging Resource Center (MIRC); electronic teaching file;
   emergency radiology; experiential; Extensible Markup Language (XML);
   web technology; computers in medicine; computer communication networks;
   data mining; database management systems; digital libraries; education;
   medical; image libraries; radiology teaching files; teaching
ID IMAGE RESOURCE CENTER; INFORMATICS; EDUCATION; UPDATE; SYSTEM
AB Teaching files are integral to radiological training. Digital Imaging
   and Communication in Medicine compatible digital radiological data and
   technological advances have made digital teaching files a desirable way
   to preserve and share representative and/or unusual cases for training
   purposes. The Medical Imaging Resource Community (MIRC) system
   developed by the Radiological Society of North America (RSNA) is a
   robust multi-platform digital teaching file implementation that is
   freely available. An emergency radiology training curriculum developed
   by the American Society of Emergency Radiology (ASER) was incorporated
   to determine if such an approach might facilitate the entry,
   maintenance, and cataloguing of interesting cases. The RSNA MIRC
   software was obtained from the main MIRC website and installed. A
   coding system was developed based on the outline form of the ASER
   curriculum. Weekly reports were generated tallying the number of cases
   in each category of the curriculum. Resident participation in the entry
   and maintenance of cases markedly increased after incorporation of the
   ASER curriculum. The coding schema facilitated progress assessment.
   Ultimately, 454 total cases were entered into the MIRC database,
   representing at least 42% of the subcategories within the ASER
   curriculum (161 out of 376). The incorporation of the ASER emergency
   radiology curriculum greatly facilitated the location, cataloguing,
   tracking, and maintenance of representative cases and served as an
   effective means by which to unify the efforts of the department to
   develop a comprehensive teaching resource within this subspecialty.
   This approach and format will be extended to other educational
   curricula in other radiological subspecialties.
```

2. Záznam časopiseckého článku ve formátu BibTeX:

```
@article{ ISI:000275551400015,
Author = {Welte, Frank J. and Kim, Sunah C. and Doshi, Devang J. and O'Connor,
Stephen C. and Coughlin, Bret F.},
Title = {{Incorporation of a Formalized Emergency Radiology Curriculum to
Facilitate Population of a MIRC-based Digital Teaching File}},
Journal = {{JOURNAL OF DIGITAL IMAGING}},
Year = {{2010}},
Volume = {{23}},
Number = {{2}},
Pages = {{226-237}},
Month = {{APR}},
Abstract = {{Teaching files are integral to radiological training. Digital Imaging
and Communication in Medicine compatible digital radiological data and
technological advances have made digital teaching files a desirable way
to preserve and share representative and/or unusual cases for training
purposes. The Medical Imaging Resource Community (MIRC) system
developed by the Radiological Society of North America (RSNA) is a
robust multi-platform digital teaching file implementation that is
freely available. An emergency radiology training curriculum developed
by the American Society of Emergency Radiology (ASER) was incorporated
to determine if such an approach might facilitate the entry,
maintenance, and cataloguing of interesting cases. The RSNA MIRC
software was obtained from the main MIRC website and installed. A
coding system was developed based on the outline form of the ASER
curriculum. Weekly reports were generated tallying the number of cases
in each category of the curriculum. Resident participation in the entry
and maintenance of cases markedly increased after incorporation of the
ASER curriculum. The coding schema facilitated progress assessment.
Ultimately, 454 total cases were entered into the MIRC database,
representing at least 42\% of the subcategories within the ASER
curriculum (161 out of 376). The incorporation of the ASER emergency
radiology curriculum greatly facilitated the location, cataloguing,
tracking, and maintenance of representative cases and served as an
effective means by which to unify the efforts of the department to
develop a comprehensive teaching resource within this subspecialty.
This approach and format will be extended to other educational
curricula in other radiological subspecialties.}},
Publisher = {{SPRINGER}},
Address = {{233 SPRING ST, NEW YORK, NY 10013 USA}},
Type = {{Article}},
Language = {{English}},
Affiliation = {{Welte, FJ (Reprint Author), Tufts Univ, Sch Med, Baystate Med Ctr,
{{}}Welte, Frank J.; Kim, Sunah C.; Doshi, Devang J.; O'Connor, Stephen C.; Coug
{{}}Coughlin, Bret F.] Hartford Hosp, Dept Radiol, Hartford, CT 06115 USA.}},
DOI = {{10.1007/s10278-009-9178-8}},
ISSN = {{0897-1889}},
Keywords = {{Medical Imaging Resource Center (MIRC); electronic teaching file;
emergency radiology; experiential; Extensible Markup Language (XML);
web technology; computers in medicine; computer communication networks;
data mining; database management systems; digital libraries; education;
medical; image libraries; radiology teaching files; teaching}},
Keywords-Plus = {{IMAGE RESOURCE CENTER; INFORMATICS; EDUCATION; UPDATE; SYSTEM}},
Subject-Category = {{Radiology, Nuclear Medicine \& Medical Imaging}},
Author-Email = {{fjwelte@gmail.com}},
Number-of-Cited-References = {{10}},
Times-Cited = {{0}},
Journal-ISO = {{J. Digit. Imaging}},
Doc-Delivery-Number = {{568VI}},
Unique-ID = {{ISI:000275551400015}},
}
```

3. Záznam časopiseckého článku ve formátu HTML:

```
FN ISI Export Format
VR 1.0
PT J
AU Welte, FJ
   Kim, SC
   Doshi, DJ
   O'Connor, SC
   Coughlin, BF
AF Welte, Frank J.
   Kim, Sunah C.
   Doshi, Devang J.
   O'Connor, Stephen C.
   Coughlin, Bret F.
TI Incorporation of a Formalized Emergency Radiology Curriculum to Facilitate Population of a
MIRC-based Digital Teaching File
SO JOURNAL OF DIGITAL IMAGING
LA English
DT Article
DE Medical Imaging Resource Center (MIRC); electronic teaching file; emergency radiology;
experiential; Extensible Markup Language (XML); web technology; computers in medicine;
computer communication networks; data mining; database management systems; digital libraries;
education; medical; image libraries; radiology teaching files; teaching
ID IMAGE RESOURCE CENTER; INFORMATICS; EDUCATION; UPDATE; SYSTEM
AB Teaching files are integral to radiological training. Digital Imaging and Communication in Medicine
compatible digital radiological data and technological advances have made digital teaching files a
desirable way to preserve and share representative and/or unusual cases for training purposes. The
Medical Imaging Resource Community (MIRC) system developed by the Radiological Society of
North America (RSNA) is a robust multi-platform digital teaching file implementation that is freely
available. An emergency radiology training curriculum developed by the American Society of
Emergency Radiology (ASER) was incorporated to determine if such an approach might facilitate
the entry, maintenance, and cataloging of interesting cases. The RSNA MIRC software was
obtained from the main MIRC website and installed. A coding system was developed based on the
outline form of the ASER curriculum. Weekly reports were generated tallying the number of cases
in each category of the curriculum. Resident participation in the entry and maintenance of cases
markedly increased after incorporation of the ASER curriculum. The coding schema facilitated
progress assessment. Ultimately, 454 total cases were entered into the MIRC database,
representing at least 42% of the subcategories within the ASER curriculum (161 out of 376). The
incorporation of the ASER emergency radiology curriculum greatly facilitated the location,
cataloging, tracking, and maintenance of representative cases and served as an effective means by
which to unify the efforts of the department to develop a comprehensive teaching resource within
this subspecialty. This approach and format will be extended to other educational curricula in other
radiological subspecialties.
C1 [Welte, Frank J.; Kim, Sunah C.; Doshi, Devang J.; O'Connor, Stephen C.; Coughlin, Bret F.]
Tufts Univ, Sch Med, Baystate Med Ctr, Dept Radiol, Springfield, MA 01199 USA.
[Coughlin, Bret F.] Hartford Hosp, Dept Radiol, Hartford, CT 06115 USA.
RP Welte, FJ, Tufts Univ, Sch Med, Baystate Med Ctr, Dept Radiol, 759 Chestnut St, Springfield,
MA 01199 USA.
EM fjwelte@gmail.com
NR 10
TC 0
PU SPRINGER
PI NEW YORK
PA 233 SPRING ST, NEW YORK, NY 10013 USA
SN 0897-1889
J9 J DIGITAL IMAGING
JI J. Digit. Imaging
PD APR
PY 2010
VL 23
IS 2
BP 226
EP 237
DI 10.1007/s10278-009-9178-8
PG 12
SC Radiology, Nuclear Medicine & Medical Imaging
GA 568VI
UT ISI:000275551400015
ER
EF
```

4. Část záznamu časopiseckého článku ve formátu čistého textu:

```

FN ISI Export Format
VR 1.0
PT J
AU Welte, FJ
    Kim, SC
    Doshi, DJ
    O'Connor, SC
    Coughlin, BF
AF Welte, Frank J.
    Kim, Sunah C.
    Doshi, Devang J.
    O'Connor, Stephen C.
    Coughlin, Bret F.
TI Incorporation of a Formalized Emergency Radiology Curriculum to
    Facilitate Population of a MIRC-based Digital Teaching File
SO JOURNAL OF DIGITAL IMAGING
LA English
DT Article
DE Medical Imaging Resource Center (MIRC); electronic teaching file;
    emergency radiology; experiential; Extensible Markup Language (XML);
    web technology; computers in medicine; computer communication networks;
    data mining; database management systems; digital libraries; education;
    medical; image libraries; radiology teaching files; teaching
ID IMAGE RESOURCE CENTER; INFORMATICS; EDUCATION; UPDATE; SYSTEM
AB Teaching files are integral to radiological training. Digital Imaging
    and Communication in Medicine compatible digital radiological data and
    technological advances have made digital teaching files a desirable way
    to preserve and share representative and/or unusual cases for training
    purposes. The Medical Imaging Resource Community (MIRC) system
    developed by the Radiological Society of North America (RSNA) is a
    robust multi-platform digital teaching file implementation that is
    freely available. An emergency radiology training curriculum developed
    by the American Society of Emergency Radiology (ASER) was incorporated
    to determine if such an approach might facilitate the entry,
    maintenance, and cataloguing of interesting cases. The RSNA MIRC
    software was obtained from the main MIRC website and installed. A
    coding system was developed based on the outline form of the ASER
    curriculum. Weekly reports were generated tallying the number of cases
    in each category of the curriculum. Resident participation in the entry
    and maintenance of cases markedly increased after incorporation of the
    ASER curriculum. The coding schema facilitated progress assessment.
    Ultimately, 454 total cases were entered into the MIRC database,
    representing at least 42% of the subcategories within the ASER
    curriculum (161 out of 376). The incorporation of the ASER emergency
    radiology curriculum greatly facilitated the location, cataloguing,
    tracking, and maintenance of representative cases and served as an
    effective means by which to unify the efforts of the department to
    develop a comprehensive teaching resource within this subspecialty.
    This approach and format will be extended to other educational
    curricula in other radiological subspecialties.
    
```

5. Část záznamu časopiseckého článku ve formátu využívajícím tabulátor jako oddělovač (určeno pro operační systém Microsoft Windows):

```

PT AU BA ED GP AF CA TI SO SE LA DT CT CY CL SP HO DE ID AB CL
J Welte, FJ; Kim, SC; Doshi, DJ; O'Connor, SC; Coughlin, BF Welte, Frank J.;
    
```

6. Část záznamu časopiseckého článku ve formátu využívajícím tabulátor jako oddělovač (určeno pro operační systém Mac OS):

```

PT AU BA ED GP AF CA TI SO SE LA DT CT CY CL SP HO DE ID AB CL
J Welte, FJ; Kim, SC; Doshi, DJ; O'Connor, SC; Coughlin, BF Welte, Frank J.;
    
```

Příloha č. 2: Scopus

1. Část záznamu časopiseckého článku v textovém formátu (ASCII):

<p>Scopus EXPORT DATE: 4 April 2010</p> <p>Wei, J. Numerical sequence matching based on local sum functions (2010) Pattern Recognition Letters, 31 (7), pp. 600-608. http://www.scopus.com/inward/record.url?eid=2-s2.0-77649342459&partnerID=40&md5=8dd</p> <p>AFFILIATIONS: Department of Computer Science, The City College of New York, 160 Con ABSTRACT: The edit distance has found a broad spectrum of applications in word proc AUTHOR KEYWORDS: Digital libraries; Dynamic programming; Edit distance; Pattern REFERENCES: Bellman, R., Dreyfus, S., (1962) Applied Dynamic Programming, , Princet Chellappa, R., Jain, A.K., (1993) Markov Random Fields, , Academic Press; Cochran, W., (1977) Sampling Techniques. third ed., , Wiley; Efron, B., Tibshirani, R., (1993) An Introduction to the Bootstrap, , Chapman and H Fukunaga, K., (1990) Statistical Pattern Recognition, , Morgan Kaufmann; Hadjidemetriou, E., Grossberg, M., Nayar, S., Multiresolution histograms and their Kolmogorov, A.N., Fomin, S.V., (1970) Introductory Real Analysis, , Dover; Ling, H., Okada, K., Diffusion distance for histogram comparison (2006) Proc. IEEE Ling, H., Okada, K., An efficient earth movers distance algorithm for robust histog Papadimitriou, C.H., Steiglitz, K., (1998) Combinatorial Optimization, , Dover; (1996) IEEE Trans. PAMI, 18 (8)., Picard R., and Pentland A. (Eds); Rubner, Y., Tomasi, C., Guibas, L., The earth movers distance as a metric for image Sakoe, H., Chiba, S., Dynamic programming algorithm optimization for spoken word re Stuart, A., Ord, K., Arnold, S., (2004) Kendalls Advanced Theory of Statistics: Cla Tenenbaum, J., de Silva, V., Langford, J., A global geometric framework for nonlinear Wei, J., Markov edit distance (2004) IEEE Trans. PAMI, 26 (3), pp. 311-321; Wei, J., Shape indexing and recognition based on regional analysis (2007) IEEE Tran Zhang, H.J., Automatic partitioning of full-motion video (1993) Multimedia Systems, CORRESPONDENCE ADDRESS: Wei, J.; Department of Computer Science, The City College o ISSN: 01678655 DOI: 10.1016/j.patrec.2009.11.009 LANGUAGE OF ORIGINAL DOCUMENT: English ABBREVIATED SOURCE TITLE: Pattern Recogn. Lett.</p>
--

2. Část záznamu časopiseckého článku ve formátu RIS (pro citační manažery Reference Manager, ProCite a EndNote):

```

TY - JOUR
T1 - Numerical sequence matching based on local sum functions
JF - Pattern Recognition Letters
JA - Pattern Recogn. Lett.
VL - 31
IS - 7
SP - 600
EP - 608
PY - 2010///
SN - 01678655 (ISSN)
AU - Wei, J.
AD - Department of Computer Science, The City College of New York, 160 Convent Ave
AB - The edit distance has found a broad spectrum of applications in word processi
KW - Digital libraries
KW - Dynamic programming
KW - Edit distance
KW - Pattern recognition
KW - Quantization schemes
N1 - Export Date: 4 April 2010
N1 - Source: Scopus
N1 - doi: 10.1016/j.patrec.2009.11.009
N1 - Language of Original Document: English
N1 - Correspondence Address: Wei, J.; Department of Computer Science, The City Col
N1 - References: Bellman, R., Dreyfus, S., (1962) Applied Dynamic Programming, , P
Chellappa, R., Jain, A.K., (1993) Markov Random Fields, , Academic Press;
Cochran, W., (1977) Sampling Techniques. third ed., , Wiley;
Efron, B., Tibshirani, R., (1993) An Introduction to the Bootstrap, , Chapman and H
Fukunaga, K., (1990) Statistical Pattern Recognition, , Morgan Kaufmann;
Hadjidemetriou, E., Grossberg, M., Nayar, S., Multiresolution histograms and their
Kolmogorov, A.N., Fomin, S.V., (1970) Introductory Real Analysis, , Dover;
Ling, H., Okada, K., Diffusion distance for histogram comparison (2006) Proc. IEEE

```

3. Část záznamu časopiseckého článku ve formátu BibTeX:

```

@Scopus
EXPORT DATE: 4 April 2010

@ARTICLE{Wei2010600,
author={Wei, J.},
title={Numerical sequence matching based on local sum functions},
journal={Pattern Recognition Letters},
year={2010},
volume={31},
number={7},
pages={600-608},
note={cited By (since 1996) 0},
url={http://www.scopus.com/inward/record.url?eid=2-s2.0-77649342459&partnerID=40&r
affiliation={Department of Computer Science, The City College of New York, 160 Con
abstract={The edit distance has found a broad spectrum of applications in word pro
author_keywords={Digital libraries; Dynamic programming; Edit distance; Pattern
references={Bellman, R., Dreyfus, S., (1962) Applied Dynamic Programming, , Prince
correspondence_address={Wei, J.; Department of Computer Science, The City College
issn={01678655},
doi={10.1016/j.patrec.2009.11.009},
language={English},
abbrev_source_title={Pattern Recogn. Lett.},
document_type={Article},
source={Scopus},
}

```


4. Část záznamu časopiseckého článku ve formátu CSV (např. pro tabulkový procesor Microsoft Excel):

```
Authors, Title, Year, Source title, Volume, Issue, Art. No., Page start, Page end, Page count  
"Wei J.", "Numerical sequence matching based on local sum functions", 2010, "Pattern Re
```

Příloha č. 3: Nature

1. Záznam časopiseckého článku ve formátu RIS:

```
TY - JOUR
AU - Zheng, Yongmei
AU - Bai, Hao
AU - Huang, Zhongbing
AU - Tian, Xuelin
AU - Nie, Fu-Qiang
AU - Zhao, Yong
AU - Zhai, Jin
AU - Jiang, Lei
TI - Directional water collection on wetted spider silk
JA - Nature
PY - 2010/02/04/print
VL - 463
IS - 7281
SP - 640
EP - 643
PB - Macmillan Publishers Limited. All rights reserved
SN - 0028-0836
UR - http://dx.doi.org/10.1038/nature08729
M3 - 10.1038/nature08729
N1 - 10.1038/nature08729
L3 - http://www.nature.com/nature/journal/v463/n7281/supinfo/nature08729\_s1.html
ER -
```

2. Záznam časopiseckého článku ve formátu citačního manažeru Connotea:

Add a bookmark

Identified: [Directional water collection on wetted spider silk](#)
Yongmei Zheng *et al.*
Nature **463** (7281), 640-3 (04 Feb 2010)
<info:doi/10.1038/nature08729>

[Edit Citation](#)

Display Title: (what you will see)

Tags: [Add to my library](#)

Tag Suggestions: (click to add)

Tags will appear here as you type in the tags box above.

Separate tags with spaces or commas. Enclose multi-word tags in "quotes". For example:
genetics "DNA structure" *history*
"C. elegans", "neuromuscular development"
See the [site guide](#) for more details.

Příloha č. 4: Science

1. Záznam časopiseckého článku ve formátu EndNote pro stejnojmenný citační manažer (pro operační systémy Microsoft Windows a Mac OS):

```
%A Mervis, Jeffrey
%T NSF Rethinks Its Digital Library
%0 Journal Article
%D 2009
%J Science
%R 10.1126/science.323.5910.54
%P 54-58
%V 323
%N 5910
%U http://www.sciencemag.org
%8 January 2, 2009
```

2. Záznam časopiseckého článku ve formátu RIS pro citační manažer Reference Manager (pro operační systém Microsoft Windows):

```
TY - JOUR
A1 - Mervis, Jeffrey
T1 - NSF Rethinks Its Digital Library
Y1 - 2009/1/2
JF - Science
JO - Science
SP - 54
EP - 58
VL - 323
IS - 5910
UR - http://www.sciencemag.org
N1 - 10.1126/science.323.5910.54
ER -
```

3. Záznam časopiseckého článku ve formátu RIS pro citační manažer ProCite (pro operační systémy Mac OS a Microsoft Windows):

```
TY - JOUR
A1 - Mervis, Jeffrey
T1 - NSF Rethinks Its Digital Library
Y1 - 2009/1/2
JF - Science
JO - Science
SP - 54
EP - 58
VL - 323
IS - 5910
UR - http://www.sciencemag.org
N1 - 10.1126/science.323.5910.54
ER -
```

4. Záznam časopiseckého článku ve formátu BibTeX (pro operační systémy Mac OS a Microsoft Windows):

```
@article{JeffreyMervis01022009,
author = {Mervis, Jeffrey},
title = {{NSF Rethinks Its Digital Library}},
journal = {Science},
volume = {323},
number = {5910},
pages = {54-58},
doi = {10.1126/science.323.5910.54},
year = {2009},
URL = {http://www.sciencemag.org},
eprint = {http://www.sciencemag.org/cgi/reprint/323/5910/54.pdf}
}
```

5. Záznam časopiseckého článku ve formátu RefWorks pro stejnojmenný citační manažer (pro operační systémy Mac OS a Microsoft Windows):

```
RT Journal
A1 Mervis, Jeffrey
T1 NSF Rethinks Its Digital Library
JF Science
JO Science
YR 2009
FD January 2
VO 323
IS 5910
SP 54
OP 58
DO 10.1126/science.323.5910.54
UL http://www.sciencemag.org
```

6. Záznam časopiseckého článku ve formátu Medlars:

```
AU - Mervis J.
TI - NSF Rethinks Its Digital Library
PT - JOURNAL ARTICLE
DP - 2009 Jan 02
TA - Science
PG - 54-58
VI - 323
IP - 5910
4099- http://www.sciencemag.org
4100- http://www.sciencemag.org/cgi/content/full/323/5910/54
SO - Science 2009 Jan 02;323(5910):54-58
```

7. Záznam časopiseckého článku ve formátu RIS pro citační manažer Zotero:

```
TY - JOUR
A1 - Mervis, Jeffrey
T1 - NSF Rethinks Its Digital Library
Y1 - 2009/1/2
JF - Science
JA - Science
SP - 54
EP - 58
VL - 323
IS - 5910
UR - http://www.sciencemag.org
M3 - 10.1126/science.323.5910.54
ER -
```

Příloha č. 5: SpringerLink

1. Část záznamu časopiseckého článku ve formátu RIS:

```
TY - JOUR
JF - International Journal on Digital Libraries
T1 - A framework for performance monitoring, load balancing, adaptive timeouts a
VL - 3
IS - 1
SP - 19
EP - 35
PY - 2000/07/25/
UR - http://dx.doi.org/10.1007/s007990050008
M3 - 10.1007/s007990050008
AU - Kapidakis, Sarantos
AU - Terzis, Sotirios
AU - Sairamesh, Jakka
N2 - In this paper, we investigate the issues of performance management in large
ER -
```

2. Část záznamu časopiseckého článku v textovém formátu:

```
Journal Title - International Journal on Digital Libraries
Article Title - A framework for performance monitoring, load balancing, adap
Volume - Volume 3
Issue - 1
First Page - 19
Last Page - 35
Issue Cover Date - 2000-07-25

Author - Sarantos Kapidakis
Author - Sotirios Terzis
Author - Jakka Sairamesh
DOI - 10.1007/s007990050008
Link - http://www.springerlink.com/content/blt6hf50nyyh30r1
```

Příloha č. 6: ScienceDirect

1. Část záznamu časopiseckého článku ve formátu RIS pro citační manažery Reference Manager, ProCite a EndNote:

```
Journal Title - International Journal on Digital Libraries
Article Title - A framework for performance monitoring, load balancing, adap
Volume - Volume 3
Issue - 1
First Page - 19
Last Page - 35
Issue Cover Date - 2000-07-25

Author - Sarantos Kapidakis
Author - Sotirios Terzis
Author - Jakka Sairamesh
DOI - 10.1007/s007990050008
Link - http://www.springerlink.com/content/blt6hf50nyyh30r1
```

2. Část záznamu časopiseckého článku v textovém formátu ASCII:

```
Mary S. Woodley, A digital library project on a shoestring, Library Collections, A
(http://www.sciencedirect.com/science/article/B6VSH-46WMTP3-3/2/91743f65924d52174b)
Abstract:
Supported by a modest grant and collaboration with local historical societies, the
Keywords: Digital library; Digital archive; Metadata standards and workflow
```

3. Část záznamu časopiseckého článku ve formátu BibTeX:

```
@article{Woodley2002199,
title = "A digital library project on a shoestring",
journal = "Library Collections, Acquisitions, and Technical Services",
volume = "26",
number = "3",
pages = "199 - 206",
year = "2002",
note = "",
issn = "1464-9055",
doi = "DOI: 10.1016/S1464-9055(02)00252-X",
url = "http://www.sciencedirect.com/science/article/B6VSH-46WMTP3-3/2/91743f65924d52174b",
author = "Mary S. Woodley",
keywords = "Digital library",
keywords = "Digital archive",
keywords = "Metadata standards and workflow",
abstract = "
Supported by a modest grant and collaboration with local historical societies, the
}
```


Příloha č. 7: Wiley InterScience

1. Část záznamu časopiseckého článku v čistém textovém formátu (určeno pro operační systém Microsoft Windows):

```
AU: Enrique Herrera-Viedma
AU: Javier López Gijón
AU: Sergio Alonso
AU: Josefina Vílchez
AU: Concha García
AU: Luis Villén
AU: Antonio Gabriel López-Herrera
TI: Applying aggregation operators for information access systems: An application
SO: International Journal of Intelligent Systems
VL: 23
NO: 12
PG: 1235-1250
YR: 2008
CP: Copyright © 2008 Wiley Periodicals, Inc., A Wiley Company
ON: 1098-111X
PN: 0884-8173
AD: Department of Computer Science and Artificial Intelligence, University of Granada
DOI: 10.1002/int.20317
US: http://dx.doi.org/10.1002/int.20317
AB: Nowadays, the information access on the Web is a main problem in the computer
```

2. Část záznamu časopiseckého článku v textovém formátu (určeno pro operační systém Mac OS):

```
AU: Enrique Herrera-Viedma
AU: Javier López Gijón
AU: Sergio Alonso
AU: Josefina Vílchez
AU: Concha García
AU: Luis Villén
AU: Antonio Gabriel López-Herrera
TI: Applying aggregation operators for information access systems: An application
SO: International Journal of Intelligent Systems
VL: 23
NO: 12
PG: 1235-1250
YR: 2008
CP: Copyright © 2008 Wiley Periodicals, Inc., A Wiley Company
ON: 1098-111X
PN: 0884-8173
AD: Department of Computer Science and Artificial Intelligence, University of Granada
DOI: 10.1002/int.20317
US: http://dx.doi.org/10.1002/int.20317
AB: Nowadays, the information access on the Web is a main problem in the computer
```

3. Část záznamu časopiseckého článku v textovém formátu (určeno pro unixové operační systémy):

```
AU: Enrique Herrera-Viedma
AU: Javier López Gijón
AU: Sergio Alonso
AU: Josefina Vílchez
AU: Concha García
AU: Luis Villén
AU: Antonio Gabriel López-Herrera
TI: Applying aggregation operators for information access systems: An application
SO: International Journal of Intelligent Systems
VL: 23
NO: 12
PG: 1235-1250
YR: 2008
CP: Copyright © 2008 Wiley Periodicals, Inc., A Wiley Company
ON: 1098-111X
EN: 0884-8173
AD: Department of Computer Science and Artificial Intelligence, University of Granada
DOI: 10.1002/int.20317
US: http://dx.doi.org/10.1002/int.20317
AB: Nowadays, the information access on the Web is a main problem in the computer
```

4. Část záznamu časopiseckého článku ve formátu EndNote (určeno pro operační systém Microsoft Windows):

```
%0 Journal Article
%A Enrique Herrera-Viedma
%A Javier López Gijón
%A Sergio Alonso
%A Josefina Vílchez
%A Concha García
%A Luis Villén
%A Antonio Gabriel López-Herrera
%T Applying aggregation operators for information access systems: An application
%J International Journal of Intelligent Systems
%V 23
%N 12
%P 1235-1250
%D 2008
%Q 1098-111X
%+ Department of Computer Science and Artificial Intelligence, University of Granada
%1 10.1002/int.20317
%U http://dx.doi.org/10.1002/int.20317
%X Nowadays, the information access on the Web is a main problem in the computer s
```

5. Část záznamu časopiseckého článku ve formátu EndNote (určeno pro operační systém Mac OS):

```
%0 Journal Article
%A Enrique Herrera-Viedma
%A Javier López Gijón
%A Sergio Alonso
%A Josefina Vilchez
%A Concha García
%A Luis Villén
%A Antonio Gabriel López-Herrera
%T Applying aggregation operators for information access systems: An application
%J International Journal of Intelligent Systems
%V 23
%N 12
%P 1235-1250
%D 2008
%@ 1098-111X
%+ Department of Computer Science and Artificial Intelligence, University of Granada
%1 10.1002/int.20317
%U http://dx.doi.org/10.1002/int.20317
%X Nowadays, the information access on the Web is a main problem in the computer
```

6. Část záznamu časopiseckého článku ve formátu EndNote (určeno pro unixové operační systémy):

```
%0 Journal Article
%A Enrique Herrera-Viedma
%A Javier López Gijón
%A Sergio Alonso
%A Josefina Vilchez
%A Concha García
%A Luis Villén
%A Antonio Gabriel López-Herrera
%T Applying aggregation operators for information access systems: An application
%J International Journal of Intelligent Systems
%V 23
%N 12
%P 1235-1250
%D 2008
%@ 1098-111X
%+ Department of Computer Science and Artificial Intelligence, University of Granada
%1 10.1002/int.20317
%U http://dx.doi.org/10.1002/int.20317
%X Nowadays, the information access on the Web is a main problem in the computer
```

Příloha č. 8: PubMed

1. Záznam časopiseckého článku ve formátu označeném jako Summary (text):

```
1: Rodríguez González D, Carpenter T, van Hemert JI, Wardlaw J. An open source toolkit for medical imaging de-identification. Eur Radiol. 2010 Mar 4. [Epub ahead of print] PubMed PMID: 20204640.
```

2. Záznam časopiseckého článku ve formátu označeném jako Abstract (text):

```
1. Eur Radiol. 2010 Mar 4. [Epub ahead of print]
```

```
An open source toolkit for medical imaging de-identification.
```

```
Rodríguez González D, Carpenter T, van Hemert JI, Wardlaw J.
```

```
National e-Science Centre, School of Informatics, University of Edinburgh, Edinburgh, UK, david.rodriguez@ed.ac.uk.
```

```
OBJECTIVE: Medical imaging acquired for clinical purposes can have several legitimate secondary uses in research projects and teaching libraries. No commonly accepted solution for anonymising these images exists because the amount of personal data that should be preserved varies case by case. Our objective is to provide a flexible mechanism for anonymising Digital Imaging and Communications in Medicine (DICOM) data that meets the requirements for deployment in multicentre trials. METHODS: We reviewed our current de-identification practices and defined the relevant use cases to extract the requirements for the de-identification process. We then used these requirements in the design and implementation of the toolkit. Finally, we tested the toolkit taking as a reference those requirements, including a multicentre deployment. RESULTS: The toolkit successfully anonymised DICOM data from various sources. Furthermore, it was shown that it could forward anonymous data to remote destinations, remove burned-in annotations, and add tracking information to the header. The toolkit also implements the DICOM standard confidentiality mechanism. CONCLUSION: A DICOM de-identification toolkit that facilitates the enforcement of privacy policies was developed. It is highly extensible, provides the necessary flexibility to account for different de-identification requirements and has a low adoption barrier for new users.
```

```
PMID: 20204640 [PubMed - as supplied by publisher]
```

3. Část záznamu časopiseckého článku ve formátu MEDLINE:

PMID- 20204640 OWN - NLM STAT- Publisher DA - 20100305 IS - 1432-1084 (Electronic) IS - 0938-7994 (Linking) DP - 2010 Mar 4 TI - An open source toolkit for medical imaging de-identification. AB - OBJECTIVE: Medical imaging acquired for clinical purposes can have several legitimate secondary uses in research projects and teaching libraries. No commonly accepted solution for anonymising these images exists because the amount of personal data that should be preserved varies case by case. Our objective is to provide a flexible mechanism for anonymising Digital Imaging and Communications in Medicine (DICOM) data that meets the requirements for deployment in multicentre trials. METHODS: We reviewed our current de-identification practices and defined the relevant use cases to extract the requirements for the de-identification process. We then used these requirements in the design and implementation of the toolkit. Finally, we tested the toolkit taking as a reference those requirements, including a multicentre deployment. RESULTS: The toolkit successfully anonymised DICOM data from various sources. Furthermore, it was shown that it could forward anonymous data to remote destinations, remove burned-in annotations, and add tracking information to the header. The toolkit also implements the DICOM standard confidentiality mechanism. CONCLUSION: A DICOM de-identification toolkit that facilitates the enforcement of privacy policies was developed. It is highly extensible, provides the necessary flexibility to account for different de-identification requirements and has a low adoption barrier for new users. AD - National e-Science Centre, School of Informatics, University of Edinburgh, Edinburgh, UK, david.rodriquez@ed.ac.uk . AU - Rodriguez Gonzalez D AU - Carpenter T AU - van Hemert JI AU - Wardlaw J	
--	--

4. Část záznamu časopiseckého článku ve formátu XML:

```
<?xml version="1.0"?>
<!DOCTYPE PubMedArticleSet PUBLIC "-//NLM//DTD PubMedArticle, 1st January 2010//EN" "http://www.ncbi.nlm.nih.gov/
<PubMedArticleSet>

<PubMedArticle>
  <MedlineCitation Status="Publisher" Owner="NLM">
    <PMID>20204640</PMID>
    <DateCreated>
      <Year>2010</Year>
      <Month>3</Month>
      <Day>5</Day>
    </DateCreated>
    <Article PubModel="Print-Electronic">
      <Journal>
        <ISSN IssnType="Electronic">1432-1084</ISSN>
        <JournalIssue CitedMedium="Internet">
          <PubDate>
            <Year>2010</Year>
            <Month>Mar</Month>
            <Day>4</Day>
          </PubDate>
        </JournalIssue>
        <Title>European radiology</Title>
      </Journal>
      <ArticleTitle>An open source toolkit for medical imaging de-identification.</ArticleTitle>
      <PageNumber>
        <MedlinePgn/>
      </PageNumber>
      <Abstract>
        <AbstractText>OBJECTIVE: Medical imaging acquired for clinical purposes can have
        several legitimate secondary uses in research projects and teaching
        libraries. No commonly accepted solution for anonymising these images exists
        because the amount of personal data that should be preserved varies case by
        case. Our objective is to provide a flexible mechanism for anonymising
        Digital Imaging and Communications in Medicine (DICOM) data that meets the
        requirements for deployment in multicentre trials. METHODS: We reviewed our
        current de-identification practices and defined the relevant use cases to
        extract the requirements for the de-identification process. We then used
        these requirements in the design and implementation of the toolkit. Finally,
        we tested the toolkit taking as a reference those requirements, including a
        multicentre deployment. RESULTS: The toolkit successfully anonymised DICOM
```

5. Část záznamu časopiseckého článku ve formátu PMID List:

```
20204640
```

Poznámka: V případě záznamu ve formátu označeném jako XML byla při zpracování změněna přípona souboru z TXT na XML, aby bylo možné obsah přehledněji zobrazit v XML editoru oXygen. Záznam ve formátu PMID List obsahuje pouze jednu položku (jeden identifikátor PMID), neboť byl proveden export pouze jednoho záznamu.

Příloha č. 9: EBSCOhost

1. Záznam časopiseckého článku v různých citačních stylech (AMA, APA, Chicago/Turabian: Author-Date, Chicago/Turabian: Humanities, MLA a Vancouver/ICMJE):

Citation Format ✕

NOTE: Review the instructions at [EBSCO Support Site](#) and make any necessary corrections before using. Pay special attention to personal names, capitalization, and dates. Always consult your library resources for the exact formatting and punctuation guidelines.

AMA (American Medical Assoc.)	Reference List Taie E, Mohamed K. The Role of Digital Libraries in Egyptian Higher Education. <i>DOMES: Digest of Middle East Studies</i> [serial online]. Fall2009 2009;18(2):40-56. Available from: Middle Eastern & Central Asian Studies, Ipswich, MA. Accessed April 5, 2010.
APA (American Psychological Assoc.)	References Taie, E., & Mohamed, K. (2009). The Role of Digital Libraries in Egyptian Higher Education. <i>DOMES: Digest of Middle East Studies</i> , 18(2), 40-56. Retrieved from Middle Eastern & Central Asian Studies database.
Chicago/Turabian: Author-Date	Reference List Taie, Eman Salman, and Khaled A. Mohamed. 2009. "The Role of Digital Libraries in Egyptian Higher Education." <i>DOMES: Digest of Middle East Studies</i> 18, no. 2: 40-56. <i>Middle Eastern & Central Asian Studies, EBSCOhost</i> (accessed April 5, 2010).
Chicago/Turabian: Humanities	Bibliography Taie, Eman Salman, and Khaled A. Mohamed. "The Role of Digital Libraries in Egyptian Higher Education." <i>DOMES: Digest of Middle East Studies</i> 18, no. 2 (Fall2009 2009): 40-56. <i>Middle Eastern & Central Asian Studies, EBSCOhost</i> (accessed April 5, 2010).
MLA (Modern Language Assoc.)	Works Cited Taie, Eman Salman, and Khaled A. Mohamed. "The Role of Digital Libraries in Egyptian Higher Education." <i>DOMES: Digest of Middle East Studies</i> 18.2 (2009): 40-56. <i>Middle Eastern & Central Asian Studies, EBSCO</i> . Web. 5 Apr. 2010.
Vancouver/ICMJE	References Taie E, Mohamed K. The Role of Digital Libraries in Egyptian Higher Education. <i>DOMES: Digest of Middle East Studies</i> [serial on the Internet]. (2009, Fall2009), [cited April 5, 2010]; 18(2): 40-56. Available from: Middle Eastern & Central Asian Studies.

[Export to Bibliographic Management Software](#) (EndNote, ProCite, Reference Manager, RefWorks, BibTeX, etc.)

[Close](#)

2. Část záznamu časopiseckého článku ve formátu pro nespécifikovaný (obecný) citační manažer:

Back
<javascript: __doPostBack('ctl100\$MainContentArea\$deliveryPrintSaveControl\$backButtons\$
1 article(s) will be saved.
To continue, in Internet Explorer, select * FILE* then * SAVE AS* from your browser's toolbar above. Be sure to save as a plain text file (.txt) or a 'Web Page, HTML only' file (.html). In Netscape, select * FILE* then * SAVE AS* from your browser's toolbar above.
Record: 1
TI- The Role of Digital Libraries in Egyptian Higher Education.
AU- Taie, Eman Salman 1
AU- Mohamed, Khaled A. 2
JN- DOMES: Digest of Middle East Studies
PD- Fall2009, Vol. 18 Issue 2, p40
PG- 17p
IL- 1 Color Photograph, 5 Charts
DT- 20090901
PT- Article
AB- Exploring the students' information, communication technology (ICT) profile and
AB- Copyright of DOMES: Digest of Middle East Studies is the property of Wiley-Black
SU- DIGITAL libraries
SU- INFORMATION & communication technologies
SU- NURSING students
SU- QUALITY assurance
SU- COLLEGE teachers
SU- EDUCATION, Higher
GE- EGYPT
AD- ^1 Helwan University
AD- ^2 El-Minia University

3. Záznam časopiseckého článku ve formátu XML:

```
<records>
<rec resultID="1">
<header shortDbName="fxh" longDbName="Middle Eastern & Central Asian Studies"
uiTerm="47395539">
<controlInfo>|
<bkinfo/>
<jinfo>
<jtl>DOMES: Digest of Middle East Studies</jtl>
<issn>10604367</issn>
</jinfo>
<pubinfo>
<dt year="2009" month="09" day="01">Fall2009</dt>
<vid>18</vid>
<iid>2</iid>
</pubinfo>
<artinfo>
<ppf>40</ppf>
<ppct>17</ppct>
<formats/>
<tig>
<att>The Role of Digital Libraries in Egyptian Higher Education.</att>
</tig>
<aug>
<au>Taie, Eman Salman</au>
<au>Mohamed, Khaled A.</au>
</aug>
<sug>
<subj type="geo">EGYPT</subj>
<subj type="naics">Libraries and Archives</subj>
<subj type="unclass">DIGITAL libraries</subj>
<subj type="unclass">INFORMATION & communication technologies</subj>
<subj type="unclass">NURSING students</subj>
<subj type="unclass">QUALITY assurance</subj>
<subj type="unclass">COLLEGE teachers</subj>
<subj type="unclass">EDUCATION, Higher</subj>
</sug>
<ab>Exploring the students' information, communication technology (ICT) profile and how it
may affect their attitudes towards the Egyptian university digital library (EUDL) and
the role of the digital library in the educational and research activities in faculties
of nursing and education in Egypt is the main concern of this study. A questionnaire was
```

4. Záznam časopiseckého článku ve formátu BibTeX:

```
Back
<javascript: __doPostBack('ctl00$MainContentArea$deliveryPrintSaveControl$backButtor

1 article(s) will be saved.

To continue, in Internet Explorer, select * FILE* then * SAVE AS* from
your browser's toolbar above. Be sure to save as a plain text file
(.txt) or a 'Web Page, HTML only' file (.html). In Netscape, select *
FILE* then * SAVE AS* from your browser's toolbar above.

EBSCO Publishing Citation Format: BibTex:
-----
References

@article{4739553920090101,
Abstract = {Exploring the students' information, communication
technology (ICT) profile and how it may affect their attitudes towards
the Egyptian university digital library (EUDL) and the role of the
digital library in the educational and research activities in faculties
of nursing and education in Egypt is the main concern of this study. A
questionnaire was distributed to students in the faculties of Nursing
and Education at five Egyptian universities. A total of 4008 responses
have been received and analyzed. The results showed that nursing
students are more active in implementing ICT and they have positive
attitudes towards the digital library. The results also revealed that
the digital library played a more positive role in the Nursing faculties
than faculties of Education. The study also revealed five barriers for
not using the digital library by the students: the restriction of
accessing EDUL as a source of information for research and education,
```

5. Záznam časopiseckého článku ve formátu MARC 21:

```
Back
<javascript: __doPostBack('ctl00$MainContentArea$deliveryPrintSaveControl$backButton$lnkBac

1 article(s) will be saved.

To continue, in Internet Explorer, select * FILE* then * SAVE AS* from
your browser's toolbar above. Be sure to save as a plain text file
(.txt) or a 'Web Page, HTML only' file (.html). In Netscape, select *
FILE* then * SAVE AS* from your browser's toolbar above.

02320naa 2200133 450000100090000008004100009016001300050100002200063700002500085700002

Back
<javascript: __doPostBack('ctl00$MainContentArea$deliveryPrintSaveControl$backButtonBottom$
```

Příloha č. 10: ProQuest Central

1. Záznam časopiseckého článku v citačním stylu APA (s nabídkou dalších stylů):

Citation style: APA, American Psychological Association (<http://www.apa.org/>)

Change citation style:

- AMA
- APA
- Chicago/Turabian/Harvard: Author-Date
- Chicago/Turabian: Humanities
- MLA
- Vancouver

Copy citation below & paste into your document

Brisaboa, N., Luaces, M., Places, N., & Seco... of documents in a geographical information retrieval system... a, 14(3), 307-331. Retrieved April 5, 2010, from ProQuest Computing. (Document ID: 1977355161).

! ATTENTION
Ensure the accuracy and completeness of your bibliography by following these [instructions](#). In particular, the formatting of author names and dates might need to be corrected and missing information may need to be filled in.

Close window

Hotovo zotero

2. Záznam časopiseckého článku ve formátu vhodném pro citační manažery ProCite, EndNote a Reference Manager (jedná se o formát RIS):

```
Provider: ProQuest LLC
Database: ProQuest
Content: application/x-research-info-systems

TY - JOUR
AU - Brisaboa,N.
AU - Luaces,M.
AU - Places,N.
AU - Seco,D.
TI - Exploiting geographic references of documents in a geographical information
JO - GeoInformatica
PY - Jul 2010
SN - 13846175
VL - 14
IS - 3
SP - 307
UR - http://proquest.umi.com/pqdweb?did=1977355161&Fmt=7&clientId=45145&RQT=309&v
AB - Issue Title: Special Issue: Semantic and Conceptual Issues in Geographic
```

3. Záznam časopiseckého článku ve formátu kompatibilním s citačními manažery ProCite, EndNote, Reference Manager a RefWorks:

```
Provider: ProQuest LLC
Database: ProQuest

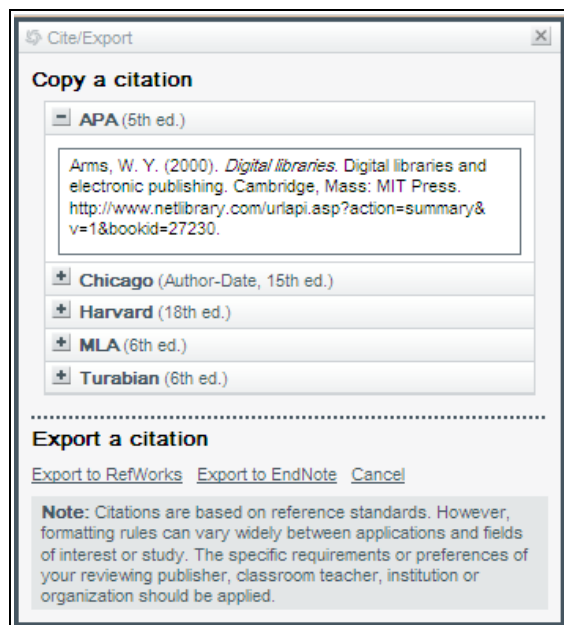
Content: text/html
DBvendor=ProQuest LLC
DBname=ProQuest
Text-encoding=ASCII

TY - JOUR

AU - Brisaboa, &nbsp;N.
AU - Luaces, &nbsp;M.
AU - Places, &nbsp;N.
AU - Seco, &nbsp;D.
TI - Exploiting geographic references of documents in a geographical information :
JO - GeoInformatica
PY - Jul 2010
SN - 13846175
VL - 14
IS - 3
SP - 307
UR - http://proquest.umi.com/pqdweb?did=1977355161&Fmt=7&clientId=45145&a
```

Příloha č. 11: WorldCat

1. Záznam časopiseckého článku v citačním stylu APA (s nabídkou dalších stylů a formátů):



2. Záznam časopiseckého článku ve formátu vhodném pro citační manažer EndNote (jedná se o formát RIS):

```
TY - ELEC
DB - /z-wcorg/
DP - http://worldcat.org
ID - 45731551
LA - English
UR - http://www.netlibrary.com/urlapi.asp?action=summary&v=1&bookid=27230
T1 - Digital libraries
A1 - Arms, William Y.
PB - MIT Press
CY - Cambridge, Mass.
Y1 - 2000///
SN - 0585281327 9780585281322
ER -
```

Příloha č. 12: bibliotek.dk

TY - BOOK
A1 - Rüger, Stefan
T1 - Multimedia information retrieval
Y1 - 2010
CY - San Rafael, Calif.
PB - Morgan & Claypool Publishers
T3 - Synthesis lectures on information concepts, retrieval, and services, 10
N1 - xi, 157 s., ill., portr., 24 cm. At its very core multimedia information retrieval means the process of searching for and finding multimedia documents; the corresponding research field is concerned with building the best possible multimedia search engines. The intriguing bit here is that the query itself can be a multimedia excerpt: For example, when you walk around in an unknown place and stumble across an interesting landmark, would it not be great if you could just take a picture with your mobile phone and send it to a service that finds a similar picture in a database and tells you more about the building - and about its significance for that matter?. Supporting users in their resource discovery mission when hunting for multimedia material is not a technological indexing problem alone. We look at interactive ways of engaging with repositories through browsing and relevance feedback, roping in geographical context, and providing visual summaries for videos. The book concludes with an overview of state-of-the-art research projects in the area of multimedia information retrieval, which gives an indication of the research and development trends and, thereby, a glimpse of the future world. This book goes further by examining the full matrix of a variety of query modes versus document types. How do you retrieve a music piece by humming? What if you want to find news video clips on forest fires using a still image? The text discusses underlying techniques and common approaches to facilitate multimedia search engines from metadata

1. Záznam časopiseckého článku ve formátu vhodném pro citační manažery Reference Manager a EndNote:

Příloha č. 13: CERN Document Server

1. Záznam časopiseckého článku ve formátu BibTeX:

```
@article{Gentil-Beccot:1214893,  
  author      = "Gentil-Beccot, A and Mele, S and Brooks, T",  
  title       = "Citing and Reading Behaviours in High-Energy Physics : How  
                a Community Stopped Worrying about Journals and Learned to  
                Love Repositories. oai:cds.cern.ch:1214893",  
  journal     = "Scientometrics",  
  number      = "arXiv:0906.5418. CERN-OPEN-2009-012. SLAC-PUB-13693",  
  pages       = "12 p",  
  month       = "Jul",  
  year        = "2009",  
}
```


2. Záznam časopiseckého článku ve formátu MARC:

```

001214893 001__ 1214893
001214893 003__ SzGeCERN
001214893 005__ 20091210090744.0
001214893 0248_ $$aoai:cds.cern.ch:1214893$$pcerncds:CERN
001214893 035__ $$9arXiv$$aoai:arXiv.org:0906.5418
001214893 035__ $$9inspire
001214893 037__ $$aarXiv:0906.5418
001214893 041__ $$aeng
001214893 088__ $$aCERN-OPEN-2009-012
001214893 088__ $$aSLAC-PUB-13693
001214893 100__ $$aGentil-Beccot, A$$uCERN
001214893 245__ $$aCiting and Reading Behaviours in High-Energy Physics : How a
Community Stopped Worrying about Journals and Learned to Love Repositories
001214893 260__ $$c2009
001214893 269__ $$aGeneva$$bCERN$$c01 Jul 2009
001214893 300__ $$a12 p
001214893 520__ $$aContemporary scholarly discourse follows many alternative
routes in addition to the three-century old tradition of publication in peer-
reviewed journals. The field of High- Energy Physics (HEP) has explored
alternative communication strategies for decades, initially via the mass mailing
of paper copies of preliminary manuscripts, then via the inception of the first
online repositories and digital libraries. This field is uniquely placed to
answer recurrent questions raised by the current trends in scholarly
communication: is there an advantage for scientists to make their work available
through repositories, often in preliminary form? Is there an advantage to
publishing in Open Access journals? Do scientists still read journals or do they
use digital repositories? The analysis of citation data demonstrates that free
and immediate online dissemination of preprints creates an immense citation
advantage in HEP, whereas publication in Open Access journals presents no
discernible advantage. In addition, the analysis of clickstreams in the leading
digital library of the field shows that HEP scientists seldom read journals,
preferring preprints instead.
001214893 595__ $$aCERN EDS
001214893 595__ $$aLANL EDS
001214893 65017 $$2SzGeCERN$$aInformation Transfer and Management
001214893 65017 $$2arXiv$$aComputing and Computers
001214893 690C_ $$aCERN
001214893 690C_ $$aARTICLE
001214893 693__ $$aNot applicable$$eNot applicable
001214893 695__ $$9LANL EDS$$acs.DL
001214893 700__ $$aMele, S$$uCERN
001214893 700__ $$aBrooks, T$$uSLAC
001214893 710__ $$5GS
001214893 773__ $$a10.1007/s11192-009-0111-1$$pScientometrics
001214893 859__ $$fanne.gentil-beccot@cern.ch
001214893 8564_ $$uhttp://arxiv.org/pdf/0906.5418.pdf$$yPreprint
001214893 8564_ $$uhttp://www.springerlink.com/content/wgx55186548v1124/$$
yJournal server
001214893 8564_ $$uhttp://cdsweb.cern.ch/record/1214893/files/CERN-OPEN-2009-
012.pdf$$yPreprint
001214893 916__ $$sn$$w200944$$ya2009
001214893 960__ $$a13
001214893 961__ $$c20100112$$h1648$$lCER01$$x20091026
001214893 963__ $$aPUBLIC
001214893 970__ $$a000704934CER
001214893 980__ $$aARTICLE
001214893 981__ $$a002829398CER01

```

3. Část záznamu časopiseckého článku ve formátu MARCXML:

```
<?xml version="1.0" encoding="UTF-8"?>
<collection xmlns="http://www.loc.gov/MARC21/slim">
  <record>
    <controlfield tag="001">1214893</controlfield>
    <controlfield tag="003">SzGeCERN</controlfield>
    <controlfield tag="005">20091210090744.0</controlfield>
    <datafield tag="035" ind1=" " ind2=" ">
      <subfield code="g">arXiv</subfield>
      <subfield code="a">oai:arXiv.org:0906.5418</subfield>
    </datafield>
    <datafield tag="035" ind1=" " ind2=" ">
      <subfield code="g">inspire</subfield>
    </datafield>
    <datafield tag="037" ind1=" " ind2=" ">
      <subfield code="a">arXiv:0906.5418</subfield>
    </datafield>
    <datafield tag="041" ind1=" " ind2=" ">
      <subfield code="a">eng</subfield>
    </datafield>
    <datafield tag="088" ind1=" " ind2=" ">
      <subfield code="a">CERN-OPEN-2009-012</subfield>
    </datafield>
    <datafield tag="088" ind1=" " ind2=" ">
      <subfield code="a">SLAC-PUB-13693</subfield>
    </datafield>
    <datafield tag="100" ind1=" " ind2=" ">
      <subfield code="a">Gentil-Beccot, A</subfield>
      <subfield code="u">CERN</subfield>
    </datafield>
    <datafield tag="245" ind1=" " ind2=" ">
      <subfield code="a">Citing and Reading Behaviours in High-Energy Physics : How a Community Stopped Worrying about Journals and Learned to Love Repositories</subfield>
    </datafield>
    <datafield tag="260" ind1=" " ind2=" ">
      <subfield code="c">2009</subfield>
    </datafield>
    <datafield tag="269" ind1=" " ind2=" ">
      <subfield code="a">Geneva</subfield>
      <subfield code="b">CERN</subfield>
      <subfield code="c">01 Jul 2009</subfield>
    </datafield>
  </record>
</collection>
```

4. Záznam časopiseckého článku ve formátu Dublin Core:

```
<?xml version="1.0" encoding="UTF-8"?>
<collection>
  <dc:dc xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
    <dc:language>eng</dc:language>
    <dc:creator>Gentil-Beccot, A</dc:creator>
    <dc:creator>Mele, S</dc:creator>
    <dc:creator>Brooks, T</dc:creator>
    <dc:title>Citing and Reading Behaviours in High-Energy Physics : How a Community Stopped
    Worrying about Journals and Learned to Love Repositories</dc:title>
    <dc:subject>Information Transfer and Management</dc:subject>
    <dc:subject>Computing and Computers</dc:subject>
    <dc:identifier>http://arxiv.org/pdf/0906.5418.pdf</dc:identifier>
    <dc:identifier>http://www.springerlink.com/content/wgx55186548v1124/</dc:identifier>
    <dc:identifier>http://cdsweb.cern.ch/record/1214893/files/CERN-OPEN-2009-012.pdf</dc:identifier>
    <dc:description>Contemporary scholarly discourse follows many alternative routes in addition to
    the three-century old tradition of publication in peer-reviewed journals. The field of High-
    Energy Physics (HEP) has explored alternative communication strategies for decades, initially
    via the mass mailing of paper copies of preliminary manuscripts, then via the inception of the
    first online repositories and digital libraries. This field is uniquely placed to answer
    recurrent questions raised by the current trends in scholarly communication: is there an
    advantage for scientists to make their work available through repositories, often in
    preliminary form? Is there an advantage to publishing in Open Access journals? Do scientists
    still read journals or do they use digital repositories? The analysis of citation data
    demonstrates that free and immediate online dissemination of preprints creates an immense
    citation advantage in HEP, whereas publication in Open Access journals presents no discernible
    advantage. In addition, the analysis of clickstreams in the leading digital library of the
    field shows that HEP scientists seldom read journals, preferring preprints instead.</dc:description>
    <dc:date>2009-10-26T17:01:46Z</dc:date>
  </dc:dc>
</collection>
```

5. Záznam časopiseckého článku ve formátu EndNote:

```
<?xml version="1.0" encoding="UTF-8"?>
<collection>
  <records>
    <record>
      <authors>
        <author>Gentil-Beccot, A</author>
        <author>Mele, S</author>
        <author>Brooks, T</author>
      </authors>
      <title>Citing and Reading Behaviours in High-Energy Physics : How a Community Stopped Worrying
        about Journals and Learned to Love Repositories</title>
      <secondary_title>Scientometrics</secondary_title>
      <pages/>
      <number/>
      <volume/>
      <abstract>Contemporary scholarly discourse follows many alternative routes in addition to the
        three-century old tradition of publication in peer-reviewed journals. The field of High-
        Energy Physics (HEP) has explored alternative communication strategies for decades,
        initially via the mass mailing of paper copies of preliminary manuscripts, then via the
        inception of the first online repositories and digital libraries. This field is uniquely
        placed to answer recurrent questions raised by the current trends in scholarly
        communication: is there an advantage for scientists to make their work available through
        repositories, often in preliminary form? Is there an advantage to publishing in Open Access
        journals? Do scientists still read journals or do they use digital repositories? The
        analysis of citation data demonstrates that free and immediate online dissemination of
        preprints creates an immense citation advantage in HEP, whereas publication in Open Access
        journals presents no discernible advantage. In addition, the analysis of clickstreams in the
        leading digital library of the field shows that HEP scientists seldom read journals,
        preferring preprints instead.</abstract>
      <keywords/>
      <year>2009</year>
      <date>2009</date>
    </record>
  </records>
</collection>
```

6. Část záznamu časopiseckého článku ve formátu NLM:

```
<?xml version="1.0" encoding="UTF-8"?>
<articles>
<article xmlns:xlink="http://www.w3.org/1999/xlink">
  <front>
    <journal-meta>
      <journal-title>Scientometrics</journal-title>
      <abbrev-journal-title>Scientometrics</abbrev-journal-title>
      <issn>0138-9130</issn>
    </journal-meta>
    <article-meta>
      <title-group>
        <article-title>Citing and Reading Behaviours in High-Energy Physics : How a Community
          Stopped Worrying about Journals and Learned to Love Repositories</article-title>
      </title-group>
      <article-id pub-id-type="doi">10.1007/s11192-009-0111-1</article-id>
      <contrib-group>
        <contrib contrib-type="author">
          <name>
            <surname>Gentil-Beccot</surname>
            <given-names>A</given-names>
          </name>
          <aff>
            <institution>CERN</institution>
          </aff>
        </contrib>
        <contrib contrib-type="author">
          <name>
            <surname>Mele</surname>
            <given-names>S</given-names>
          </name>
          <aff>
            <institution>CERN</institution>
          </aff>
        </contrib>
        <contrib contrib-type="author">
          <name>
            <surname>Brooks</surname>
            <given-names>T</given-names>
          </name>
        </contrib>
      </contrib-group>
    </article-meta>
  </front>
</article>
</articles>
```

Příloha č. 14: E-LIS

1. Záznam časopiseckého článku ve formátu HTML Citation:

```
Cassella, Maria and Calvi, Licia ECDL 2009: enhancing digital libraries users' experience. D-Lib Magazine, 2009, vol. 15, n. 11/12. [Journal Article (On-line/Unpaginated)]
```

2. Zdrojový kód záznamu časopiseckého článku ve formátu HTML Citation:

```
<p class="citation"><span class="person_name">Cassella, Maria</span> and <span class="person_name">Calvi, Licia</span> <a href="http://eprints.rcslis.org/17305/">ECDL 2009: enhancing digital libraries users' experience.</a> <i>D-Lib Magazine</i>, 2009, vol. 15, n. 11/12. [Journal Article (On-line/Unpaginated)]</p>
```

3. Záznam časopiseckého článku ve formátu ASCII Citation:

```
Cassella, Maria and Calvi, Licia ECDL 2009: enhancing digital libraries users' experience. D-Lib Magazine, 2009, vol. 15, n. 11/12. [Journal Article (On-line/Unpaginated)]
```

4. Záznam časopiseckého článku ve formátu Full Metadata:

Full Metadata	
creators_name:	Cassella, Maria
creators_name:	Calvi, Licia
type:	journal
datestamp:	2009-11-27 08:27:45
lastmod:	2009-11-27 08:27:45
status_changed:	2009-11-27 08:27:45
userid:	2170
metadata_visibility:	show
countries:	IT-
title:	ECDL 2009: enhancing digital libraries users' experience
ispublished:	pub
subjects:	H.
subjects:	L.
full_text_status:	public
keywords:	Digital libraries, Knowledge organization, Resource Discovery, Information Retrieval, Preservation, Usability
abstract:	Report of the 13th European Conference on Digital Libraries (ECDL 2009), Corfù, 27 September-2 October 2009
othersabs_othabstracts:	Report della tredicesima European Conference on Digital Libraries (ECDL 2009), Corfù, 27 September-2 October 2009
date:	2009-11-15
publication:	D-Lib Magazine
volume:	15
number:	11/12
publisher:	D-Lib Alliance
linguabib:	en
refereed:	FALSE
altloc:	http://www.dlib.org/dlib/november09/calvi/11calvi.html
document_url:	http://eprints.rclis.org/17305/1/11calvi.html
lang_full_codex:	en

5. Zdrojový kód záznamu časopiseckého článku ve formátu Full Metadata:

```
<html><head></head><body><h3>Full Metadata</h3><table border=2>
<tr><td>creators_name: </td><td>Cassella, Maria</td></tr>
<tr><td>creators_name: </td><td>Calvi, Licia</td></tr>
<tr><td>type: </td><td>journale</td></tr>
<tr><td>datestamp: </td><td>2009-11-27 08:27:45</td></tr>
<tr><td>lastmod: </td><td>2009-11-27 08:27:45</td></tr>
<tr><td>status_changed: </td><td>2009-11-27 08:27:45</td></tr>
<tr><td>userid: </td><td>2170</td></tr>
<tr><td>metadata_visibility: </td><td>show</td></tr>
<tr><td>countries: </td><td>IT-</td></tr>
<tr><td>title: </td><td>ECDL 2009: enhancing digital libraries users'
experience</td></tr>
<tr><td>ispublished: </td><td>pub</td></tr>
<tr><td>subjects: </td><td>H.</td></tr>
<tr><td>subjects: </td><td>L.</td></tr>
<tr><td>full_text_status: </td><td>public</td></tr>
<tr><td>keywords: </td><td>Digital libraries, Knowledge organization, Resource
Discovery, Information Retrieval, Preservation, Usability</td></tr>
<tr><td>abstract: </td><td>Report of the 13th European Conference on Digital
Libraries (ECDL 2009), Corfù, 27 September-2 October 2009 </td></tr>
<tr><td>othersabs_othabstracts: </td><td>Report della tredicesima European
Conference on Digital Libraries (ECDL 2009), Corfù, 27 September-2 October
2009 </td></tr>
<tr><td>date: </td><td>2009-11-15</td></tr>
<tr><td>publication: </td><td>D-Lib Magazine</td></tr>
<tr><td>volume: </td><td>15</td></tr>
<tr><td>number: </td><td>11/12</td></tr>
<tr><td>publisher: </td><td>D-Lib Alliance</td></tr>
<tr><td>linguabib: </td><td>en</td></tr>
<tr><td>refereed: </td><td>FALSE</td></tr>
<tr><td>altloc: </td><td>http://www.dlib.org/dlib/november09/calvi
/11calvi.html</td></tr>
<tr><td>document_url: </td><td>http://eprints.rcis.org/17305
/1/11calvi.html</td></tr>
<tr><td>lang_full_codex: </td><td>en</td></tr>
</table></body></html>
```


6. Záznam časopiseckého článku ve formátu OpenURL ContextObject:

```
<ctx:context-object xmlns:xsi="http://www.w3.org/2001/XMLSchema" xmlns:ctx="info:ofi/fmt:xml:xd:ctx"
timestamp="2009-11-27T08:27:45Z"
xsi:schemaLocation="info:ofi/fmt:xml:xd:ctx http://www.openurl.info/registry/docs/info:ofi/fmt:xml:xd:ctx">
<ctx:referent>
<ctx:identifier>info:oai:eprints.rclis.org:17305</ctx:identifier>
<ctx:metadata-by-val>
<ctx:format>info:ofi/fmt:xml:xd:oai_dc</ctx:format>
<ctx:metadata>
<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<dc:title>ECDL 2009: enhancing digital libraries users' experience</dc:title>
<dc:creator>Cassella, Maria</dc:creator>
<dc:creator>Calvi, Licia</dc:creator>
<dc:subject>H. Information sources, supports, channels.</dc:subject>
<dc:subject>L. Information technology and library technology.</dc:subject>
<dc:description>Report of the 13th European Conference on Digital Libraries
(ECDL 2009), Corfù, 27 September-2 October 2009 </dc:description>
<dc:publisher>D-Lib Alliance</dc:publisher>
<dc:date>2009-11-15</dc:date>
<dc:type>Journal Article (On-line/Unpaginated)</dc:type>
<dc:type>NonPeerReviewed</dc:type>
<dc:format>text/html</dc:format>
<dc:relation>http://www.dlib.org/dlib/november09/calvi/11calvi.html</dc:relation>
<dc:identifier>http://eprints.rclis.org/17305/</dc:identifier>
<dc:language>en</dc:language>
</oai_dc:dc>
</ctx:metadata>
</ctx:metadata-by-val>
</ctx:referent>
</ctx:context-object>
```

7. Záznam časopiseckého článku ve formátu EndNote:

```
%0 Generic
%A Cassella, Maria
%A Calvi, Licia
%D 2009
%F eprints3:17305
%I D-Lib Alliance
%K Digital libraries, Knowledge organization, Resource Discovery, Information
Retrieval, Preservation, Usability
%N 11/12
%T ECDL 2009: enhancing digital libraries users' experience
%U http://eprints.rclis.org/17305/
%V 15
%X Report of the 13th European Conference on Digital Libraries (ECDL 2009),
Corfù, 27 September-2 October 2009
```

8. Záznam časopiseckého článku ve formátu BibTeX:

```
@misc{eprints317305,  
  volume = {15},  
  number = {11/12},  
  month = {November},  
  author = {Maria Cassella and Licia Calvi},  
  title = {ECDL 2009: enhancing digital libraries users' experience},  
  publisher = {D-Lib Alliance},  
  journal = {D-Lib Magazine},  
  year = {2009},  
  keywords = {Digital libraries, Knowledge organization, Resource  
Discovery, Information Retrieval, Preservation, Usability},  
  url = {http://eprints.rclis.org/17305/},  
  abstract = {Report of the 13th European Conference on Digital Libraries  
(ECDL 2009), Corf?, 27 September-2 October 2009 }  
}
```

9. Záznam časopiseckého článku ve formátu OpenURL ContextObject in Span:

```
?url_ver=Z39.88-2004&rft_val_fmt=info%3Aofi%2Ffmt%3Akev%3Amtx%3Adc&rft.title=  
ECDL+2009%3A+enhancing+digital+libraries+users'+experience&rft.creator=Cassella%  
2C+Maria&rft.creator=Calvi%2C+Licia&rft.subject=H.+Information+sources%2C+  
supports%2C+channels.&rft.subject=L.+Information+technology+and+library+  
technology.&rft.description=Report+of+the+13th+European+Conference+on+Digital+  
Libraries+(ECDL+2009)%2C+Corf%3%83%2%B9%2C+27+September-2+October+2009&rft.  
publisher=D-Lib+Alliance&rft.date=2009-11-15&rft.type=Journal+Article+(On-line%  
2FUnpaginated)&rft.type=NonPeerReviewed&rft.format=text%2Fhtml&rft.relation=  
http%3A%2F%2Fwww.dlib.org%2Fdlib%2Fnovember09%2Fcalvi%2F11calvi.html&rft.  
identifier=http%3A%2F%2Feprints.rclis.org%2F17305%2F&rft.language=en
```

10. Záznam časopiseckého článku ve formátu MODS:

```
<?xml:mods xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:mods="http://www.loc.gov/mods/v3" version="3.0"
xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
  <mods:titleInfo>
    <mods:title>ECDL 2009: enhancing digital libraries users' experience</mods:title>
  </mods:titleInfo>
  <mods:name type="personal">
    <mods:namePart type="given">Maria</mods:namePart>
    <mods:namePart type="family">Cassella</mods:namePart>
    <mods:role>
      <mods:roleTerm type="text">author</mods:roleTerm>
    </mods:role>
  </mods:name>
  <mods:name type="personal">
    <mods:namePart type="given">Licia</mods:namePart>
    <mods:namePart type="family">Calvi</mods:namePart>
    <mods:role>
      <mods:roleTerm type="text">author</mods:roleTerm>
    </mods:role>
  </mods:name>
  <mods:abstract>Report of the 13th European Conference on Digital Libraries (ECDL 2009), Corfù,
  27 September-2 October 2009 </mods:abstract>
  <mods:classification authority="lcc">H. Information sources, supports, channels.</mods:classification>
  <mods:classification authority="lcc">L. Information technology and library technology.</mods:classification>
  <mods:originInfo>
    <mods:dateIssued encoding="iso8061">2009-11-15</mods:dateIssued>
  </mods:originInfo>
  <mods:originInfo>
    <mods:publisher>D-Lib Alliance</mods:publisher>
  </mods:originInfo>
  <mods:genre>Journal Article (On-line/Unpaginated)</mods:genre>
</mods:mods>
```

11. Část záznamu časopiseckého článku ve formátu DIDL:

```
<did:DIIDL xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:didl="urn:mpeg:mpeg21:2002:02-DIDL-NS"
xsi:schemaLocation="urn:mpeg:mpeg21:2002:02-DIDL-NS
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG-21_schema_files/did/didmodel.xsd">
<didl:Item>
<didl:Description>
<didl:Statement mimeType="application/xml; charset=utf-8">
<dii:Identifier xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg21:2002:01-DII-NS
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG-21_schema_files/dii/dii.xsd"
xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS"
>http://eprints.rclis.org/17305/</dii:Identifier>
</didl:Statement>
</didl:Description>
<didl:Description>
<didl:Statement mimeType="application/xml; charset=utf-8">
<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<dc:title>ECDL 2009: enhancing digital libraries users' experience</dc:title>
<dc:creator>Cassella, Maria</dc:creator>
<dc:creator>Calvi, Licia</dc:creator>
<dc:subject>H. Information sources, supports, channels.</dc:subject>
<dc:subject>L. Information technology and library technology.</dc:subject>
<dc:description>Report of the 13th European Conference on Digital Libraries
(ECDL 2009), Corfú, 27 September-2 October 2009 </dc:description>
<dc:publisher>D-Lib Alliance</dc:publisher>
<dc:date>2009-11-15</dc:date>
<dc:type>Journal Article (On-line/Unpaginated)</dc:type>
<dc:type>NonPeerReviewed</dc:type>
<dc:format>text/html</dc:format>
<dc:relation>http://www.dlib.org/dlib/november09/calvi/11calvi.html</dc:relation>
<dc:identifier>http://eprints.rclis.org/17305/</dc:identifier>
<dc:language>en</dc:language>
</oai_dc:dc>
</didl:Statement>
```

12. Část záznamu časopiseckého článku ve formátu EP3 XML:

```
<eprint id="http://eprints.rclis.org/id/eprint/17305" xmlns="http://eprints.org/ep2/data/2.0">
  <eprintid>17305</eprintid>
  <rev_number>15</rev_number>
  <eprint_status>archive</eprint_status>
  <userid>2170</userid>
  <dir>disk0/00/01/73/05</dir>
  <datestamp>2009-11-27 08:27:45</datestamp>
  <lastmod>2009-11-27 08:27:45</lastmod>
  <status_changed>2009-11-27 08:27:45</status_changed>
  <type>journale</type>
  <metadata_visibility>show</metadata_visibility>
  <item_issues_count>0</item_issues_count>
  <abstract>Report of the 13th European Conference on Digital Libraries (ECDL 2009), Corfù, 27
  September-2 October 2009</abstract>
  <othersabs>
    <item>
      <othabstracts>Report della tredicesima European Conference on Digital Libraries (ECDL 2009),
      Corfù, 27 September-2 October 2009</othabstracts>
      <language>it</language>
    </item>
  </othersabs>
  <allloc>
    <item>http://www.dlib.org/dlib/november09/calvi/11calvi.html</item>
  </allloc>
  <creators>
    <item>
      <name>
        <family>Cassella</family>
        <given>Maria</given>
      </name>
      <id/>
    </item>
    <item>
      <name>
        <family>Calvi</family>
        <given>Licia</given>
      </name>
      <id/>
    </item>
  </creators>
```

13. Záznam časopiseckého článku ve formátu ISO Citation:

Cassella, Maria and Calvi, Licia [ECDL 2009: enhancing digital libraries users' experience](http://eprints.rclis.org/17305). *D-Lib Magazine*, 2009, vol. 15, n. 11/12 [cited 2/4/2010]. Available from WWW: <http://eprints.rclis.org/17305>

14. Zdrojový kód záznamu časopiseckého článku ve formátu ISO Citation:

```
<p class="citation"><span class="person_name">Cassella, Maria</span> and <span class="person_name">Calvi, Licia</span> <a href="http://eprints.rclis.org/17305/">ECDL 2009: enhancing digital libraries users' experience.</a> <i>D-Lib Magazine</i>, 2009, vol. 15, n. 11/12.</p></div><script type="text/javascript">
// 
var currentTime = new Date();var month = currentTime.getMonth() + 1;var day =
currentTime.getDate();var year = currentTime.getFullYear();document.write("[cited
" + month + "/" + day + "/" + year + ".]")
// ]]&gt;&lt;/script&gt; Available from WWW:
  &lt;a rel="self" href="http://eprints.rclis.org/17305"&gt;http://eprints.rclis.org
/17305&lt;/a&gt;&lt;/p&gt;</pre></div><div data-bbox="115 278 599 296" data-label="Section-Header"><h4>15. Záznam časopiseckého článku ve formátu Dublin Core:</h4></div><div data-bbox="118 314 806 507" data-label="Text"><pre>title: ECDL 2009: enhancing digital libraries users' experience
creator: Cassella, Maria
creator: Calvi, Licia
subject: H. Information sources, supports, channels.
subject: L. Information technology and library technology.
description: Report of the 13th European Conference on Digital Libraries (ECDL
2009), Corfù, 27 September-2 October 2009
publisher: D-Lib Alliance
date: 2009-11-15
type: Journal Article (On-line/Unpaginated)
type: NonPeerReviewed
format: text/html
relation: <a href="http://www.dlib.org/dlib/november09/calvi/11calvi.html">http://www.dlib.org/dlib/november09/calvi/11calvi.html</a>
identifier: <a href="http://eprints.rclis.org/17305/">http://eprints.rclis.org/17305/</a>
language: en</pre></div><div data-bbox="115 530 657 547" data-label="Section-Header"><h4>16. Záznam časopiseckého článku ve formátu Reference Manager:</h4></div><div data-bbox="118 565 806 797" data-label="Text"><pre>TY - GEN
ID - eprints317305
UR - <a href="http://www.dlib.org/dlib/november09/calvi/11calvi.html">http://www.dlib.org/dlib/november09/calvi/11calvi.html</a>
A1 - Cassella, Maria
A1 - Calvi, Licia
Y1 - 2009/11/15/
N2 - Report of the 13th European Conference on Digital Libraries (ECDL 2009),
Corfù, 27 September-2 October 2009
PB - D-Lib Alliance
KW - Digital libraries
KW - Knowledge organization
KW - Resource Discovery
KW - Information Retrieval
KW - Preservation
KW - Usability
TI - EC DL 2009: enhancing digital libraries users' experience
AV - public
ER -</pre></div><div data-bbox="186 895 213 912" data-label="Page-Footer"><p>44</p></div>
```

17. Část záznamu časopiseckého článku ve formátu Eprints Application Profile:

```
creators_name: Cassella, Maria
creators_name: Calvi, Licia
type: journale
datestamp: 2009-11-27 08:27:45
lastmod: 2009-11-27 08:27:45
metadata_visibility: show
title: ECDL 2009: enhancing digital libraries users' experience
ispublished: pub
subjects: H.
subjects: L.
full_text_status: public
keywords: Digital libraries, Knowledge organization, Resource Discovery,
Information Retrieval, Preservation, Usability
abstract: Report of the 13th European Conference on Digital Libraries (ECDL
2009), Corfù, 27 September-2 October 2009
date: 2009-11-15
publication: D-Lib Magazine
volume: 15
number: 11/12
publisher: D-Lib Alliance
refereed: FALSE
altloc: http://www.dlib.org/dlib/november09/calvi/11calvi.html
citation: Cassella, Maria and Calvi, Licia ECDL 2009: enhancing digital
libraries users' experience. D-Lib Magazine, 2009, vol. 15, n. 11/12. [Journal
Article (On-line/Unpaginated)]
document_url: http://eprints.rclis.org/17305/1/11calvi.html
```

18. Záznam časopiseckého článku ve formátu Simple Metadata:

```
<epdxc:descriptionSet xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:epdxc="http://purl.org/eprint/epdxc/2006-11-16/"
xsi:schemaLocation="http://purl.org/eprint/epdxc/2006-11-16/ http://purl.org/eprint/epdxc/xsd/2006-11-16/epdxc.xsd">
<epdxc:description epdxc:resourceURI="http://eprints.rclis.org/17305/">
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/type"
epdxc:valueURI="http://purl.org/eprint/entitType/ScholarlyWork"/>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/identifier">
<epdxc:valueString epdxc:sesURI="http://purl.org/dc/terms/URI"
>http://eprints.rclis.org/17305/</epdxc:valueString>
</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/eprint/terms/isExpressedAs"
epdxc:valueRef="expression_17305_1"/>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/title">
<epdxc:valueString>ECDL 2009: enhancing digital libraries users'
experience</epdxc:valueString>
</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/dc/terms/abstract">
<epdxc:valueString>Report of the 13th European Conference on Digital Libraries (ECDL 2009),
Corfu, 27 September-2 October 2009</epdxc:valueString>
</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/dc/elements/1.1/creator"
>Cassella, Maria</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/dc/elements/1.1/creator">Calvi,
Licia</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/subject">
<epdxc:valueString>H.</epdxc:valueString>
</epdxc:statement>
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/subject">
<epdxc:valueString>L.</epdxc:valueString>
</epdxc:statement>
</epdxc:description>
<epdxc:description epdxc:resourceId="expression_17305_1">
<epdxc:statement epdxc:propertyURI="http://purl.org/dc/elements/1.1/title">
<epdxc:valueString>ECDL 2009: enhancing digital libraries users'
experience</epdxc:valueString>
</epdxc:statement>
```

19. Záznam časopiseckého článku ve formátu Refer:

```
%A Maria Cassella
%A Licia Calvi
%J D-Lib Magazine
%T ECDL 2009: enhancing digital libraries users' experience
%X Report of the 13th European Conference on Digital Libraries (ECDL 2009),
Corfu?, 27 September-2 October 2009
%N 11/12
%K Digital libraries, Knowledge organization, Resource Discovery, Information
Retrieval, Preservation, Usability
%V 15
%D 2009
%I D-Lib Alliance
%L eprints317305
```


20. Záznam časopiseckého článku ve formátu METS:

```
<mets:mets OBJID="oai:eprints.rclis.org:17305" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:mods="http://www.loc.gov/mods/v3" LABEL="Eprints Item"
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:mets="http://www.loc.gov/METS/">
<mets:metsHdr CREATEDATA="2010-02-04T17:29:04Z">
  <mets:agent TYPE="ORGANIZATION" ROLE="CUSTODIAN">
    <mets:name>E-LIS</mets:name>
  </mets:agent>
</mets:metsHdr>
<mets:dmdSec ID="DMD_oai:eprints.rclis.org:17305_mods">
  <mets:mdWrap MDTYPE="mods">
    <mets:xmlData>
      <mods:titleInfo>
        <mods:title>ECDL 2009: enhancing digital libraries users'
experience</mods:title>
      </mods:titleInfo>
      <mods:name type="personal">
        <mods:namePart type="given">Maria</mods:namePart>
        <mods:namePart type="family">Cassella</mods:namePart>
        <mods:role>
          <mods:roleTerm type="text">author</mods:roleTerm>
        </mods:role>
      </mods:name>
      <mods:name type="personal">
        <mods:namePart type="given">Licia</mods:namePart>
        <mods:namePart type="family">Calvi</mods:namePart>
        <mods:role>
          <mods:roleTerm type="text">author</mods:roleTerm>
        </mods:role>
      </mods:name>
      <mods:abstract>Report of the 13th European Conference on Digital Libraries (ECDL
2009), Corfù, 27 September-2 October 2009 </mods:abstract>
      <mods:classification authority="lcc">H. Information sources, supports, channels.</mods:classification>
      <mods:classification authority="lcc">L. Information technology and library
technology.</mods:classification>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>
</mets:mets>
```

Příloha č. 15: GOLX-3-47

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="GOLX-3-47">
      <head></head>
      <body>
        <type>journal-article</type>
        <title>Nomenclature of tetrapyrroles</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry and International Union of Biochemistry, Joint
            Commission on Biochemical Nomenclature</author>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
        </creators>
        <position type="page" subtype="start">779</position>
        <position type="page" subtype="end">832</position>
        <identifier type="doi">doi:10.1351/pac198759060779</identifier>
        <identifier type="url">http://iupac.org/publications/pac/59/6/0779/</identifier>
        <source document-id="GOLX-3-47a"/>
      </body>
    </document>

    <document id="GOLX-3-47a">
      <head></head>
      <body>
        <type>issue</type>
        <title type="short">6</title>
        <time>1987</time>
        <source document-id="GOLX-3-47b"/>
      </body>
    </document>

    <document id="GOLX-3-47b">
      <head></head>
```

```
<body>
  <type>volume</type>
  <title type="short">59</title>
  <time>1987</time>
  <source document-id="GOLX-3-47c"/>
</body>
</document>

<document id="GOLX-3-47c">
  <head></head>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>

</reference>
```

Příloha č. 16: GOLX-3-58

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="GOLX-3-58">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Nomenclature of steroids</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry and International Union of Biochemistry, Joint
            Commission on Biochemical Nomenclature</author>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
        </creators>
        <position type="page" subtype="start">1783</position>
        <position type="page" subtype="end">1822</position>
        <identifier type="doi">doi:10.1351/pac198961101783</identifier>
        <identifier type="url">http://iupac.org/publications/pac/61/10/1783/</identifier>
        <source document-id="GOLX-3-58a"/>
      </body>
    </document>

    <document id="GOLX-3-58a">
      <head/>
      <body>
        <type>issue</type>
        <title type="short">10</title>
        <time>1989</time>
        <source document-id="GOLX-3-58b"/>
      </body>
    </document>

    <document id="GOLX-3-58b">
      <head/>
```

```
<body>
  <type>volume</type>
  <title type="short">61</title>
  <time>1989</time>
  <source document-id="GOLX-3-58c"/>
</body>
</document>

<document id="GOLX-3-58c">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>

</reference>
```

Příloha č. 17: GOLX-3-93

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="GOLX-3-93">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Glossary of class names of organic compounds and reactivity intermediates
          based on structure</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry, Organic Chemistry Division, Commission on
            Nomenclature of Organic Chemistry (III.1), Commission on Physical
            Organic Chemistry (III.2)</author>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
          <contributor type="edition">
            <first-name type="initial">P.</first-name>
            <middle-name type="initial">A.</middle-name>
            <middle-name type="initial">S.</middle-name>
            <surname>Smith</surname>
          </contributor>
          <contributor type="edition">
            <first-name type="initial">D.</first-name>
            <surname>Tavernier</surname>
          </contributor>
        </creators>
        <position type="page" subtype="start">1307</position>
        <position type="page" subtype="end">1375</position>
        <identifier type="doi">doi:10.1351/pac199567081307</identifier>
        <identifier type="url">http://iupac.org/publications/pac/67/8/1307/</identifier>
        <source document-id="GOLX-3-93a"/>
      </body>
    </document>
  </reference>
```

```
<document id="GOLX-3-93a">
  <head/>
  <body>
    <type>issue</type>
    <title type="short">8-9</title>
    <time>1995</time>
    <source document-id="GOLX-3-93b"/>
  </body>
</document>

<document id="GOLX-3-93b">
  <head/>
  <body>
    <type>volume</type>
    <title type="short">67</title>
    <time>1995</time>
    <source document-id="GOLX-3-93c"/>
  </body>
</document>

<document id="GOLX-3-93c">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>

</reference>
```

Příloha č. 18: GOLX-3-103

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="GOLX-3-103">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Basic terminology of stereochemistry</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry, Organic Chemistry Division, Commission on
            Nomenclature of Organic Chemistry (III.1), Commission on Physical
            Organic Chemistry (III.2)</author>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
        </creators>
        <position type="page" subtype="start">2193</position>
        <position type="page" subtype="end">2222</position>
        <identifier type="doi">doi:10.1351/pac199668122193</identifier>
        <identifier type="url">http://iupac.org/publications/pac/68/12/2193/</identifier>
        <source document-id="GOLX-3-103a"/>
      </body>
    </document>

    <document id="GOLX-3-103a">
      <head/>
      <body>
        <type>issue</type>
        <title type="short">12</title>
        <time>1996</time>
        <source document-id="GOLX-3-103b"/>
      </body>
    </document>

    <document id="GOLX-3-103b">
```



```
<head/>
<body>
  <type>volume</type>
  <title type="short">68</title>
  <time>1996</time>
  <source document-id="GOLX-3-103c"/>
</body>
</document>

<document id="GOLX-3-103c">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>

</reference>
```

Příloha č. 19: RED3-13-13A

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="RED3-13-13A">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Nomenclature of tetrapyrroles</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry and International Union of Biochemistry, Joint
            Commission on Biochemical Nomenclature</author>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
        </creators>
        <position type="page" subtype="start">779</position>
        <position type="page" subtype="end">832</position>
        <identifier type="doi">doi:10.1351/pac198759060779</identifier>
        <identifier type="url">http://iupac.org/publications/pac/59/6/0779/</identifier>
        <source document-id="RED3-13-13Aa"/>
      </body>
    </document>

    <document id="RED3-13-13Aa">
      <head/>
      <body>
        <type>issue</type>
        <title type="short">6</title>
        <time>1987</time>
        <source document-id="RED3-13-13Ab"/>
      </body>
    </document>

    <document id="RED3-13-13Ab">
      <head/>
```

```
<body>
  <type>volume</type>
  <title type="short">59</title>
  <time>1987</time>
  <source document-id="RED3-13-13Ac"/>
</body>
</document>

<document id="RED3-13-13Ac">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>

</reference>
```

Příloha č. 20: RED3-14-7D

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="RED3-14-7D">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Corrections to 'A Guide to IUPAC Nomenclature of Organic
        Compounds'</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
            Applied Chemistry, Organic Chemistry Division, Commission on
            Nomenclature of Organic Chemistry</author>
          <contributor type="edition">
            <first-name>Henri</first-name>
            <middle-name type="initial">A.</middle-name>
            <surname>Favre</surname>
          </contributor>
          <contributor type="edition">
            <first-name>Karl-Heinz</first-name>
            <surname>Hellwich</surname>
          </contributor>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
          <contributor type="edition">
            <first-name>Warren</first-name>
            <middle-name type="initial">H.</middle-name>
            <surname>Powell</surname>
          </contributor>
          <contributor type="edition">
            <first-name>James</first-name>
            <middle-name type="initial">G.</middle-name>
            <surname>Traynham</surname>
          </contributor>
        </creators>
      </body>
    </document>
  </reference>
```

```

</creators>
<position type="page" subtype="start">1327</position>
<position type="page" subtype="end">1330</position>
<identifier type="doi">doi:10.1351/pac199971071327</identifier>
<identifier type="url">http://iupac.org/publications/pac/71/7/1327/</identifier>
<source document-id="RED3-14-7Da"/>
</body>
</document>

<document id="RED3-14-7Da">
  <head/>
  <body>
    <type>issue</type>
    <title type="short">7</title>
    <time>1999</time>
    <source document-id="RED3-14-7Db"/>
  </body>
</document>

<document id="RED3-14-7Db">
  <head/>
  <body>
    <type>volume</type>
    <title type="short">71</title>
    <time>1999</time>
    <source document-id="RED3-14-7Dc"/>
  </body>
</document>

<document id="RED3-14-7Dc">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>
</reference>

```

Příloha č. 21: RED3-45-1C

```
<?xml version="1.0" encoding="UTF-8"?>
  <reference>
    <document id="RED3-45-1C">
      <head/>
      <body>
        <type>journal-article</type>
        <title>Corrections to 'A Guide to IUPAC Nomenclature of Organic
        Compounds'</title>
        <creators>
          <author type="group" subtype="institution">International Union of Pure and
          Applied Chemistry, Organic Chemistry Division, Commission on
          Nomenclature of Organic Chemistry</author>
          <contributor type="edition">
            <first-name>Henry</first-name>
            <middle-name type="initial">A.</middle-name>
            <surname>Favre</surname>
          </contributor>
          <contributor type="edition">
            <first-name>Karl-Heinz</first-name>
            <surname>Hellwich</surname>
          </contributor>
          <contributor type="edition">
            <first-name type="initial">G.</first-name>
            <middle-name type="initial">P.</middle-name>
            <surname>Moss</surname>
          </contributor>
          <contributor type="edition">
            <first-name>Warren</first-name>
            <middle-name type="initial">H.</middle-name>
            <surname>Powell</surname>
          </contributor>
          <contributor type="edition">
            <first-name>James</first-name>
            <middle-name type="initial">G.</middle-name>
            <surname>Traynham</surname>
          </contributor>
        </creators>
      </body>
    </document>
  </reference>
```

```
</creators>
<position type="page" subtype="start">1327</position>
<position type="page" subtype="end">1330</position>
<identifier type="doi">doi:10.1351/pac199971071327</identifier>
<identifier type="url">http://iupac.org/publications/pac/71/7/1327/</identifier>
<source document-id="RED3-45-1Ca"/>
</body>
</document>

<document id="RED3-45-1Ca">
  <head/>
  <body>
    <type>issue</type>
    <title type="short">7</title>
    <time>1999</time>
    <source document-id="RED3-45-1Cb"/>
  </body>
</document>

<document id="RED3-45-1Cb">
  <head/>
  <body>
    <type>volume</type>
    <title type="short">71</title>
    <time>1999</time>
    <source document-id="RED3-45-1Cc"/>
  </body>
</document>

<document id="RED3-45-1Cc">
  <head/>
  <body>
    <type>journal</type>
    <title>Pure and Applied Chemistry</title>
    <title convention="cassi">Pure Appl. Chem.</title>
    <title convention="iupac">PAC</title>
  </body>
</document>
</reference>
```