

Česká zemědělská univerzita v Praze
Provozně ekonomická fakulta
Katedra informačních technologií



Metadatové formáty v kontextu efektivního sdílení obohacených dat

Disertační práce

Autor: Ing. Michal Stočes
Školitel: doc. Ing. Zdeněk Havlíček, CSc.

2017

Poděkování

Děkuji svým školitelům, panu doc. Ing. Zdeňku Havlíčkovi, CSc. za odborné vedení, pomoc a podněty při zpracování tématu práce a paní doc. PhDr. Ivaně Švarcové, CSc., in memoriam, za cenné rady na začátku mého doktorského studia.

Dále pak děkuji svým kolegům z Katedry informačních technologií za připomínky a rady.

Děkuji za cenné podněty při formulování pracovních hypotéz kolegům z řešitelského týmu mezinárodního projektu VOA3R (Virtual Open Access Agriculture & Aquaculture Repository Project).

V neposlední řadě pak děkuji své rodině a obzvláště mé manželce Petře, která mi je životní oporou a inspirací. Bez její laskavé péče o mě a dcery Alici a Žofii by tato práce nemohla vzniknout.

Metadatové formáty v kontextu efektivního sdílení obohacených dat

Souhrn

Hlavním cílem předkládané disertační práce je analyzovat současný stav webových aplikací pro podporu vědecké činnosti a navrhnout a ověřit metodiku pro zkvalitnění sdílení dat mezi těmito aplikacemi. Práce je rozdělena na dvě části. V teoretické části práce je literární rešerše zaměřená na problematiku současné úrovně webových aplikací pro podporu vědecké činnosti, a to z těchto hledisek metadatových formátů popisu digitálních artefaktů (objektů) a jejich sdílení pomocí webových služeb, otevřeného obsahu a stavu a funkcí aplikací typu vědecká sociální síť.

V praktické části práce je navržen nový metodický postup pro zkvalitnění přenosu metadat, které jsou obohaceny (data enrichment) činnostmi aplikací typu sociální síť pro vědecké pracovníky (Social network service for scientists). Navržený metodický postup je následně ověřen vytvořením aplikačního profilu aplikace a záznamu konkrétního digitálního artefaktu. V závěrečné části jsou shrnuty přínosy nové metodiky a naznačeny další možné směry výzkumu v této oblasti.

Klíčová slova

metadata, aplikační profil, sociální sítě pro vědecké pracovníky, sdílení dat

Metadata formats for efficient sharing of enriched data

Abstract

The main goal of this dissertation is to analyze the current state of web applications in the field of scientific support and to design and validate methodology for improving data sharing between these applications. The thesis is divided into two parts. The theoretical part of the thesis consists of a literary research focused on web applications for support of scientific activity currently in use. It also contains an analysis of related aspects such as metadata formats for description of digital artifacts (objects), metadata sharing through web services, open content and functions of scientific social networks.

In the practical part of the thesis, a new methodical approach for metadata transfer is proposed. Its main purpose is to improve the transfer of metadata which is enriched by social network services for scientists. The proposed method is then verified by creating an application profile and a recording of a specific digital artifact. The final part of the thesis summarizes the benefits of the new methodology and outlines other possible directions of research in this field.

Keywords

metadata, application profile, social network service for scientists, data sharing

Obsah

1	ÚVOD	8
2	CÍLE PRÁCE	10
3	METODICKÝ POSTUP	11
4	TEORETICKÁ VÝCHODISKA.....	16
4.1	VYMEZENÍ POJMŮ.....	16
4.1.1	<i>Data.....</i>	16
4.1.2	<i>Informace.....</i>	16
4.1.3	<i>Znalosti</i>	17
4.1.4	<i>Shrnutí</i>	17
4.2	METADATOVÉ FORMÁTY A JEJICH SDÍLENÍ	18
4.2.1	<i>Způsob zápisu metadat.....</i>	18
4.2.2	<i>Aplikační profil.....</i>	28
4.2.3	<i>Řízené slovníky deskriptorů.....</i>	32
4.2.4	<i>Webové služby</i>	37
4.2.5	<i>Technologie webových služeb.....</i>	38
4.2.6	<i>Technologie pro sdílení dat.....</i>	47
4.2.7	<i>Shrnutí</i>	51
4.3	MOŽNOSTI ZVEŘEJNĚNÍ OBSAHU	53
4.3.1	<i>Otevřená data (open data)</i>	53
4.3.2	<i>Licence</i>	55
4.3.3	<i>Otevřený přístup</i>	56
4.3.4	<i>Shrnutí</i>	57
4.4	APLIKACE PRO PODPORU VĚDY.....	58
4.4.1	<i>Digitální knihovny.....</i>	58
4.4.2	<i>Aplikace využívající metadata.....</i>	60
4.4.3	<i>Shrnutí</i>	64
4.5	POSKYTOVÁNÍ METADAT PRO PODPORU VĚDY	65
4.5.1	<i>Analýza a zhodnocení vybraných institucionálních repositářů</i>	65
4.5.2	<i>Identifikace digitálního artefaktu</i>	66
4.5.3	<i>Identifikace osob.....</i>	68
4.5.4	<i>Analýza sociálních sítí pro vědecké pracovníky.....</i>	69

4.5.5	<i>Shrnutí</i>	70
4.6	METADATOVÉ MODELY V PRAXI	72
4.6.1	<i>Vazba a identifikace autora</i>	72
4.6.2	<i>Vazba a identifikace reference, citace</i>	73
4.6.3	<i>Mapování metadatových modelů</i>	74
4.6.4	<i>Shrnutí</i>	76
4.7	HYPOTÉZA DISERTAČNÍ PRÁCE	77
5	NÁVRH NOVÉ METODIKY	78
5.1	PŘEDPOKLADY PRO POUŽITÍ METODIKY	78
5.2	FORMULACE METODIKY MSOD	78
5.2.1	<i>Etapa 1 Identifikace metadat</i>	78
5.2.2	<i>Etapa 2 Tvorba aplikačního profilu</i>	79
5.3	OČEKÁVANÉ PŘÍNOSY NAVRŽENÉ METODIKY	83
6	OVĚŘENÍ A ZHODNOCENÍ METODIKY MSOD	85
6.1	SESTAVENÍ APLIKAČNÍHO PROFILU	85
6.1.1	<i>Etapa 1 Identifikace metadat</i>	85
6.1.2	<i>Etapa 2 Tvorba aplikačního profilu</i>	85
6.2	TVORBA ZÁZNAMU DIGITÁLNÍHO ARTEFAKTU (OBJEKTU)	88
6.3	OMEZENÍ PRO POUŽITÍ NAVRHOVANÉ METODIKY	89
6.4	ZHODNOCENÍ A PŘÍNOSY NAVRHOVANÉ METODIKY	89
7	ZÁVĚRY DISERTAČNÍ PRÁCE	90
7.1	SPLNĚNÍ DÍLČÍCH CÍŮ DISERTAČNÍ PRÁCE	90
7.2	PŘÍNOSY DISERTAČNÍ PRÁCE PRO VĚDU A PRAXI	96
7.3	DISKUSE A NÁMĚTY K DALŠÍMU VĚDECKÉMU VÝZKUMU	96
8	SEZNAM LITERATURY	97
9	SEZNAM PŘÍLOH	104
A.	SEZNAM OBRÁZKŮ	105
B.	SEZNAM TABULEK	106
C.	SEZNAM GRAFŮ	107
D.	SEZNAM POŽITÝCH ZKRATEK	108
E.	DEFINICE POUŽITÝCH POJMŮ	110
F.	ANALÝZA INSTITUCIONÁLNÍCH REPOSITÁŘŮ	111

G.	SCHÉMA APLIKAČNÍHO PROFILU.....	115
H.	XML ZÁZNAM	117
I.	PUBLIKAČNÍ ČINNOST AUTORA	119

1 Úvod

„Metadata liberates us, liberates knowledge.“

David Weinberger
2008, IT žurnalista

Rozvoj informačních a komunikačních technologií velmi přispěl ke zrychlení procesů. Příkladem může být publikace zpráv. V dobách před masivním rozšířením moderních technologií musel autor článku svůj rukopis zaslat vydavateli, ten článek předal sazeči, který připravil texty pro tisk. Po vytištění bylo nutno materiály doručit čtenářům. Proces trval v optimálních podmínkách v řádech hodin až několika dní. V současné době stačí autorovi jednoduše kliknout myší a z pohodlí své pracovny okamžitě zveřejnit článek na Internetu. Takto rychle publikované informace si může prostřednictvím počítačové sítě přečíst kdokoliv téměř kdekoliv na světě, a to ihned po vydání. Uživatelé Internetu sami vytvářejí jeho obsah. Zrychlení a jednoduchost procesů má samozřejmě i svá negativa. Internet je přesycen různými druhy informací. Informační síť vytváří džungli, ve které je těžké se vyznat. Chybně označené záznamy se v informační síti ztratí. Dohledatelnost konkrétních informací závisí na tom, jak moc dobře jsou digitální objekty zkatalogizovány a sdíleny.

Čtenáři musí být ostražití při výběru relevantních informací. S růstem množství informací klesá kvalita celé sítě. Internet je plný článků, šířících polopravdy, nedůsledně vystavených na relevantních datech. Pro budoucí rozvoj informačních sítí je důležitá podpora kvality informací. Nepotřebujeme kvantitu ale kvalitu! Problematika přemíry informací zasáhla všechny oblasti lidských činností. Předkládaná disertační práce se zaměřuje na oblast publikování a sdílení vědeckých informací.

Fenoménem posledních dob, který výrazně změnil komunikaci v lidské společnosti, jsou sociální sítě. Tato nová komunikační platforma zasáhla i oblast vědy a výzkumu. Každá vědecká práce začíná důsledným prostudováním současného stavu dané problematiky. K podpoře této

činnosti vznikla v poslední době řada specializovaných převážně webových aplikací. Příkladem těchto aplikací je např. sociální síť pro vědecké pracovníky (Social network service for scientists) ResearchGate¹ nebo oborově zaměřená sociální síť VOA3R² (Virtual Open Access Agriculture & Aquaculture Repository Project).

Většina vědeckých pracovníků je nucena pracovat s více aplikacemi. V každé aplikaci je uživatel vyzván k vytvoření profilu a nahrání metadat vědeckých publikací. Každá aplikace vytvoří uživateli a jeho publikacím specifický identifikátor. Je obtížné takto vytvořená obohacující data v systému exportovat nebo přenášet data z jedné aplikace do druhé. Předkládaná disertační práce „Metadatové formáty v kontextu efektivního sdílení obohacených dat“ si klade za cíl na základě porovnání a studia různých aplikací navrhnout metodický postup pro zkvalitnění přenosu těchto dat a tím i poskytnout jednodušší práci s vědeckými informacemi.

¹ <https://www.researchgate.net/>

² <http://voa3r.redpanda.gr/>

2 Cíle práce

Hlavním cílem předkládané disertační práce je analyzovat současný stav webových aplikací pro podporu vědecké činnosti a navrhnout a ověřit metodiku pro zkvalitnění sdílení dat mezi těmito aplikacemi.

Dílčí cíle disertační práce jsou:

1. Prozkoumat současnou úroveň webových aplikací pro podporu vědecké činnosti, a to z těchto hledisek:
 - a. metadatových formátů popisu digitálních artefaktů (objektů) a jejich sdílení pomocí webových služeb,
 - b. otevřeného obsahu,
 - c. stavu a funkcí aplikací typu vědecká sociální síť.
2. Analyzovat vybrané institucionální repositáře a způsoby poskytování dat a identifikaci subjektů v aplikacích typu sociální síť pro vědecké pracovníky (Social network service for scientists).
3. Zhodnotit problematiku a možnosti zápisu identifikace autorů, referencí a citací vybranými metadatovými modely.
4. Navrhnout postup pro zkvalitnění práce s daty vzniklými činnostmi aplikací typu sociální síť pro vědecké pracovníky.
5. Ověřit správnost postupu vytvořením aplikačního profilu digitálního artefaktu (objektu).
6. Vytvořit záznam vybraného digitálního artefaktu na základě pravidel vytvořeného aplikačního profilu.
7. Formulovat závěry včetně zhodnocení navrženého postupu.

3 Metodický postup

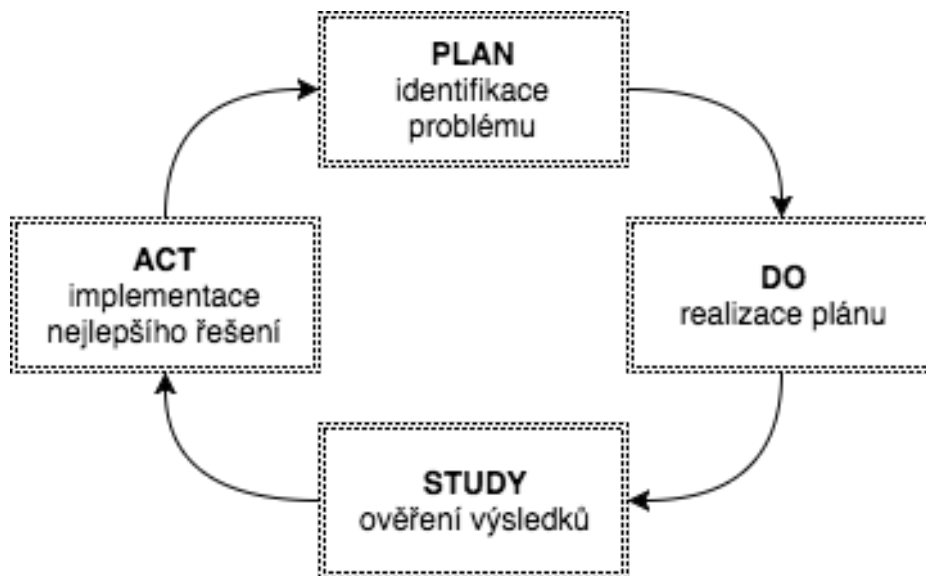
Předkládaná disertační práce „Metadatové formáty v kontextu efektivního sdílení obohacených dat“ je zaměřena na vybraná témata, definovaná cílem práce a skládá se z následujících kroků:

- **V prvním kroku** práce je analyzována problematika na základě literární rešerše zahraniční a české vědecké literatury, odborných vědeckých časopisů a příspěvků z konferencí. Jsou upřednostňovány recenzované texty, zaindexované v předních vědeckých citačních databázích Scopus a Web of Science (WoS). V závěru kapitoly jsou formulována zhodnocení jako syntéza získaných poznatků. Dále na základě analýzy jsou vybrány vhodné metadatové modely a formáty dat pro popis digitálních artefaktů.
- **Ve druhém kroku** je analyzována problematika otevřenosti, metadatové popisy a jejich sdílení vybraných českých institucionálních repositářů.
- **Ve třetím kroku** jsou rozebrány způsoby, metody a techniky poskytování dat a identifikace subjektů v aplikacích typu sociální síť pro vědecké pracovníky.
- **Ve čtvrtém kroku** jsou analyzovány vybrané metadatové modely z pohledu, jakým způsobem implementují problematiku identifikace autora, digitálního artefaktu, citací a referencí.
- **V pátém kroku** práce je navržen metodický postup pro tvorbu aplikačního profilu obohacujících dat digitálního artefaktu (objektu).
- **V šestém kroku** je navržený postup ověřen vytvořením aplikačního profilu pro popis obohacených dat digitálních artefaktů (objektů) aplikací typu sociální síť pro vědecké pracovníky.
- **V sedmém kroku** je vytvořen metadatový záznam konkrétního vybraného digitálního artefaktu (objektu).
- **V osmém kroku** jsou zhodnoceny výsledky a formulovány obecné závěry jako syntéza získaných poznatků. Dále jsou zde představeny náměty na výzkum v dané oblasti.

K dosažení cílů práce byly použity základní vědecké metody - analýza, syntéza, indukce, komparace a analogie. S ohledem na cíl práce, zkvalitnění sdílení dat, jsou aplikovány dvě metody používané v procesu zlepšování.

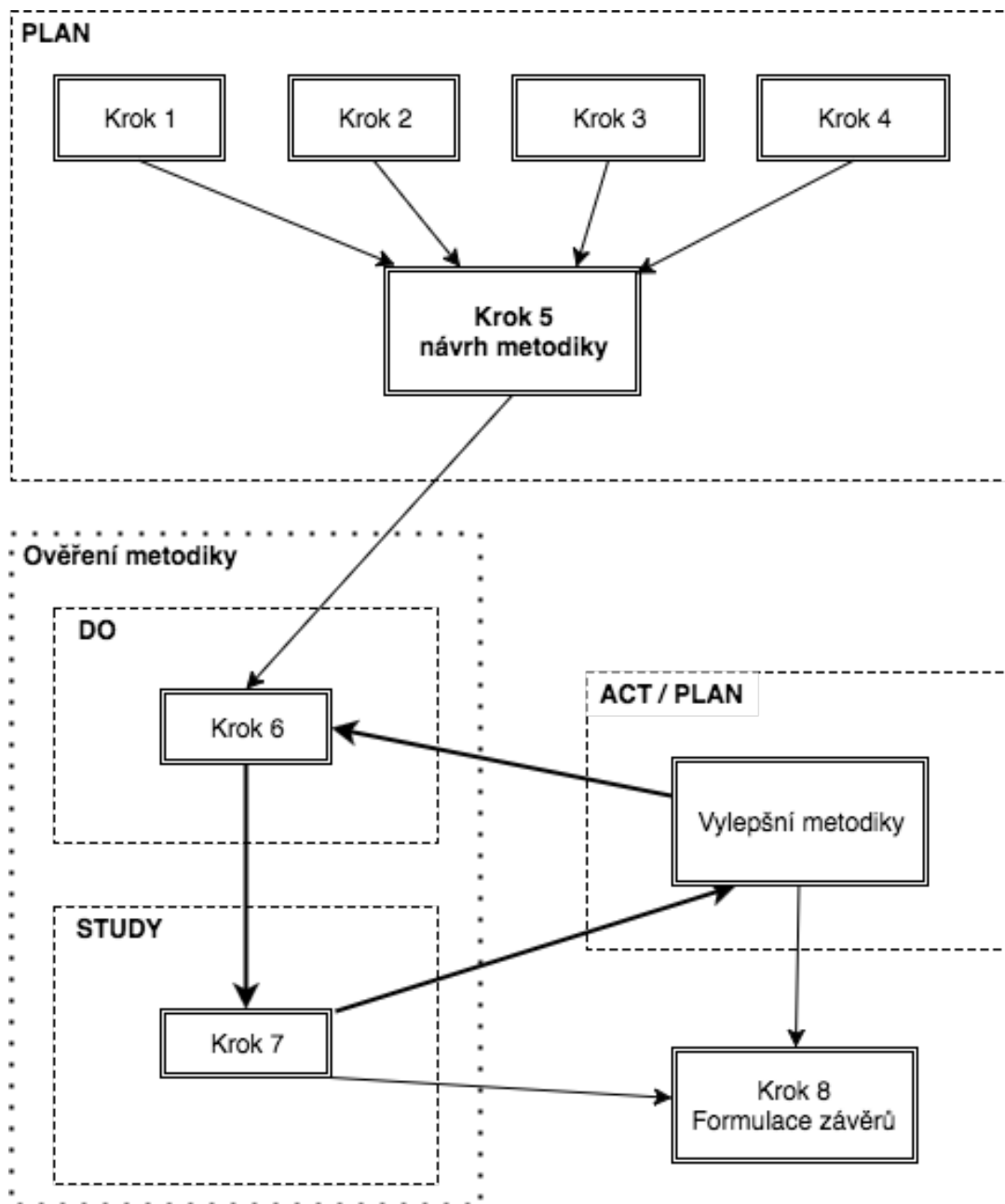
První metoda je popsána cyklem PDSA, známá také jako Demingův cyklus (Deming Wheel). Vychází z anglického modelu plan-do-study-act (plánuj-udělej-studuj-uskutečni). Jedná se o metodu postupného zlepšování v mnoha oborech, včetně informačních technologií. Skládá se z následujících fází: (Rao, a další, 1996), (Doucek, a další, 2013), (Deming, 2016):

- fáze 1 – P (plan) – identifikace problémů (záměr),
- fáze 2 – D (do) – realizace plánu,
- fáze 3 – S (study) – ověření výsledku realizace oproti původnímu záměru,
- fáze 4 – A (act) – úpravy záměru i vlastního provedení na základě ověření a implementace zlepšení do praxe, implementace nejlepšího řešení.



Obrázek 1 – Metoda PDSA. Zdroj: Autor.

Použití cyklu PDSA při návrhu postupu je znázorněn na obrázku 2.



Obrázek 2 - PDLA ve vztahu s kroky postupu vypracování disertační práce. Zdroj: Autor.

V souvislosti s cyklem PDSA bývá často také zmiňována druhá metoda – metoda sedmi kroků (seven-step method). (Rao, a další, 1996):

Fáze PDSA	Metoda sedmi kroků
PLAN (naplánuj)	Identifikace problému a jeho jasné vymezení
	Analýza aktuálního stavu
	Identifikace možných příčin problému
DO (proved')	Plánování a implementace řešení
STUDY (studuj)	Zhodnocení výsledků
ACT (jednej)	Standardizace řešení
	Zhodnocení navrženého řešení a návrh budoucích plánů/opatření

Tabulka 1 – Vztah PDSA k metodě sedmi kroků (Rao, a další, 1996)

Metodický aparát:

Metodologie

Metodologie (z řec. *methodos*, sledování, stopování, od *hodos*, cesta) je vědní disciplína, která se zabývá metodami, jejich tvorbou a aplikací. Pod vlivem angličtiny se slovo někdy používá i pro metodiku (použitý postup či metodu).³

Metoda

Metoda (z řeckého *met-hodos* – doslova „za cestou“, cesta za něčím) je postup nebo návod, jak získávat správné poznatky a prostředek poznání. Moderní ideál metody jako obecného návodu, jak získávat spolehlivé poznatky, na něž se druzí mohou spolehnout (Descartes, 1992).

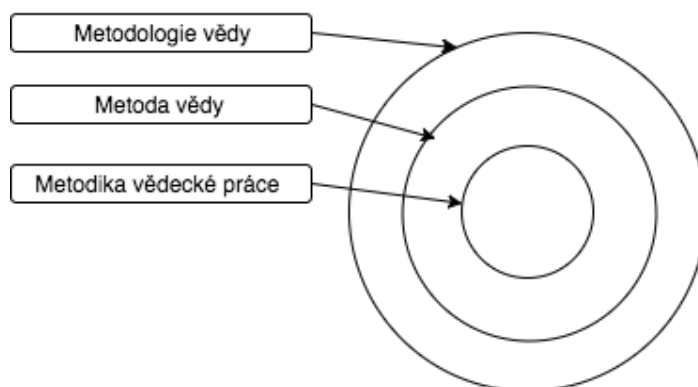
Metodika

Dle Širokého (2011) je metodika definována jako „konkrétní postup řešení určitého problému, vymezuje přesně jednotlivé kroky postupu nebo

³ <http://www.dictionary.com/browse/methodology>

konkrétní metody. Metodika je určitý návod, jak dosáhnout výzkumného cíle, může být schematicky znázorněna vývojovým diagramem či podobným formalizovaným zápisem. Metodiku lze obecně označit jako pracovní postup nebo návod“.

Závislosti mezi „metodologií vědy“, „metodou vědy“ a „metodikou vědecké práce“ dle Ochrany (2010) zachycuje diagram na obrázku 3.



Obrázek 3 - Závislosti mezi „metodologií vědy“, „metodou vědy“ a „metodikou vědecké práce“. Zdroj: Autor dle (Ochrana, 2010)

4 Teoretická východiska

Rozbor problematiky předkládané disertační práce volně vychází z obhájené rigorózní práce autora. Převážná část původní analýzy současného stavu poznání rigorózní práce byla aktualizována dle nejnovějších poznatků a přepracována vzhledem ke stanoveným cílům této práce. Poznatky získané literární rešerší jsou v následujících kapitolách utříděny do celkového přehledu problematiky a v závěru rešeršní části práce jsou formulována zhodnocení jako syntéza získaných poznatků.

4.1 Vymezení pojmů

4.1.1 Data

Pojem data je označením pro text, čísla, vizualizace, zvuk, popř. další smyslové vjemy, které existují v podobě posloupnosti znaků či signálů. Dle normy ISO/IEC 24765 (ISO/IEC, 2008), *Systems and Software Engineering Vocabulary* jsou data vyjádřením faktů, instrukcí nebo způsobem konceptů vhodným pro komunikaci, interpretaci nebo zpracování lidmi nebo počítači. Z hlediska zpracování dat v aplikacích jsou data formou vyjádření informace zpracovávanou aplikací a jeho uživateli.

4.1.2 Informace

Teorie informace zkoumá matematickou reprezentaci podmínek a parametrů, ovlivňujících přenos a zpracování informací.

Definice teorie informace dle Vaňkové (2006) „Teorie informace je klíčovou disciplínou pro jakoukoli oblast informatiky. Popisuje, jakým způsobem můžeme pracovat s informacemi, jakým způsobem je možné je kódovat, jaká jsou omezení v přenosu dat. Jde tedy o fundamentální vědu, bez jejíchž – alespoň základních – znalostí se v oblasti teoretické informatiky nebo počítačových sítí můžeme jen velmi obtížně pohnout dále.“

Při procesu archivace dat je především nutné data zabezpečit tak, aby jejich uložení bylo dlouhodobé. Dalším důležitým aspektem archivace

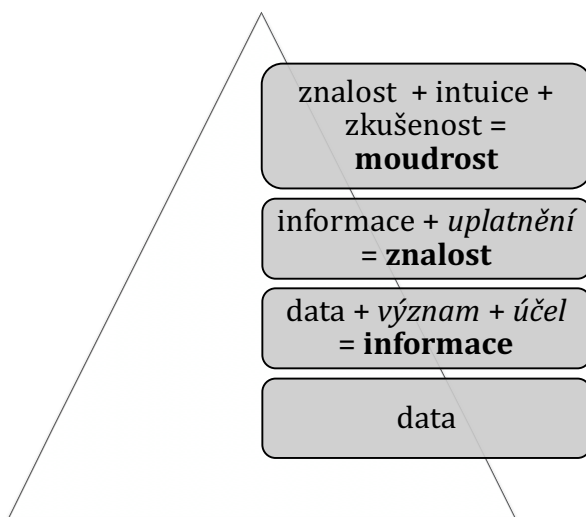
dat je jejich správná katalogizace. Absence dokumentace o datech nebo jejich chybná katalogizace může způsobit, že data jsou nedohledatelná a tím pádem se s nimi velmi těžko dále pracuje. (Garofalo, 2015) (Cejpek, 2012)

4.1.3 Znalosti

Vodáček a Rosický (1997) definují znalosti jako „proměnný systém se vzájemnou interakcí zkušeností, faktů, hodnot, myšlenkových procesů a významů. Vytvářejí též rámec pro vyhodnocování a integraci nových informací. Znalosti jsou výsledkem aktivního učení se. Znalosti jsou využívány k identifikaci a interpretaci dat.“ Z manažerského pohledu na znalosti lze uvést slova Petera Druckera (1995): „Informace jsou data, obohacená o relevantnost a účelnost, přeměna dat v informace tudíž vyžaduje znalosti“.

4.1.4 Shrnutí

Vztah závislosti pojmů data, informace, znalost a moudrost je znázorněn v pyramidovém diagramu viz Obrázek 4.



Obrázek 4 – Vztah pojmů data, informace, znalost. Zdroj: Autor

4.2 Metadatové formáty a jejich sdílení

4.2.1 Způsob zápisu metadat

Slovo „metadata“ bylo poprvé použito P. R. Bagleyem (1968) v roce 1968 v knize „Extension of programming language concepts“.

Metadata se nejčastěji a zjednodušeně definují jako „data o datech“ nebo „informace o informacích“, ale existuje celá řada dalších a komplexnějších definic. Definice dle Branda (2003): „Metadata jsou strukturovaná data, která popisují, vysvětlují, lokalizují a umožňují jednodušší získání, používání a spravování informačního zdroje.“, definice dle Berana (2013) říká: „Metadata jsou strukturovaná data – parametry – poskytující základní popisné informace o digitálních objektech, jejichž primárním účelem je usnadnění vyhledávání ve zdrojích. Zahrnují elementární informace o primárních datech, strukturované dle určitých pravidel a standardů, čímž zefektivňují spravování velkého množství objektů v datových strukturách“. Jedna z mnoha formálních definic uvádí, že metadata jsou data sdružená s objekty, která zbavují potenciální uživatele nutnosti předběžné znalosti existence či charakteristik těchto objektů. (Bartošek, 1999)

Samotná metadata mohou být dále popisována dalšími metadaty. Metadata jsou psána dle stanovených pravidel, a tudíž jsou strojově čitelná. (Beran, 2013).

Metadata se dají rozdělit do následujících skupin dle Brethertona, (1994):

- *Popisná (deskriptivní) metadata* slouží k jednoznačné identifikaci dokumentu. Obsahují například název, jméno autora, klíčová slova apod.
- *Strukturální metadata* ukazují, z jakých částí se dokument skládá dohromady, například číslování stránek, kapitol apod.
- *Administrativní metadata* obsahují technické informace o dokumentu (formát), přístupová práva apod.

Metadata slouží především k vyhledávání a záležitostem s vyhledáváním spojených. Mezi vybrané funkce metadat patří dle Bartoška (1999):

- *funkce dokumentační* (popis důležitých charakteristik informačního zdroje),
- *funkce identifikační* (jednoznačná identifikace nezávislá na čase a prostoru),
- *funkce vyhledávací* (zjištění existence zdroje), lokalizační (kde je zdroj umístěn),
- *funkce selekční* (výběr zdrojů na základě jejich jmenných či věcných charakteristik).

Vybrané metadatové formáty

Existuje celá řada metadatových formátů používaných v knihovnictví, Bratková (1999) uvádí např. Text Encoding Initiative (*TEI*) Independent Headers, Encoding Archival Description (*EAD*), Computer Interchange of Museum Information (*CIMI*), Global Information Locator Service (*GILS*), Machine-Readable Cataloging (*MARC*), Dublin Core (DC), Efix (Exchangeable image file format), *Metadata Object Description Schema (MODS)* a další.

Další skupina metadatových formátů slouží pro popis různých druhů digitálních objektů. Příkladem může být formát EXIF (Exchangeable image file format), sloužící pro vkládání popisů digitálních fotografií, nebo kontejner metadat ID3 tag, kterým lze přidat dodatečné informace k digitálnímu audio záznamu. (Chang, 2005)

4.2.1.1 Standard Dublin Core

Pro standard Dublin Core je v odborné literatuře používána zkratka DC. DC je soubor patnácti metadatových prvků a jeho hlavním účelem je usnadnit vyhledání elektronických zdrojů. DC byl vyvinut odborníky z různých oblastí (počítačová věda, knihovnictví). Soubor prvků DC je

standardizován dle norem ISO 15836:2009, (Mezinárodní organizace pro normalizaci) poslední aktualizace 2014 (ISO, 2009) a ANCI/NISO Z39.85 (American National Standards Institute / National Information Standards Organization) z roku 2007. (ANCI/NISO, 2007) DC je v současné době udržován organizací *Dublin Core Metadata Initiative*. (DCMI) (DCMI, 2016) Všechny z patnácti metadatových prvků jsou nepovinné. Seznam prvků DC se nachází v tabulce 2.

<ul style="list-style-type: none"> • Název (Title) • Tvůrce (Creator) • Předmět (Subject) • Popis (Description) • Vydavatel (Publisher) • Příspěvatel (Contributor) • Datum (Date) • Typ (Type) 	<ul style="list-style-type: none"> • Formát (Format) • Identifikátor (Identifier) • Zdroj (Source) • Jazyk (Language) • Vztah (Relation) • Pokrytí (Coverage) • Práva (Rights)
---	---

Tabulka 2 – DC - základní prvky. Zdroj: (Dublin Core : Czech, 2006)

Prvotní účel DC spočíval ve funkci popisování digitálních dokumentů zveřejněných na internetu přímo autorem. Pro svůj univerzální návrh začal být využíván institucemi, zabývajícími se formálním zpracováním zdrojů (muzea, knihovny, univerzity apod.). DC lze použít k popisu digitálních i nedigitálních objektů. (Dublin Core : Czech, 2006)

Kvalifikovaný a nekvalifikovaný Dublin Core

DC je dle způsobu zápisu hodnot k prvkům rozdělován do dvou typů, tzv. *Jednoduchý Dublin Core* nebo také *Nekvalifikovaný Dublin Core* (Simple Dublin Core, Unqualified Dublin Core) a tzv. *Kvalifikovaný Dublin Core*. U *jednoduchého Dublin Core* hodnoty prvků nejsou nijak omezeny oproti *Kvalifikovanému Dublin Core*, kde jsou upřesněna omezení pro hodnoty prvků

pomocí *kvalifikovaných termínů* (qualified terms) a *kvalifikátorů*. Formáty pro vkládání dat vycházejí z všeobecně uznávaných standardů. (Hodge, 2005)

Formáty zápisu Dublin Core a jeho prvky

Metadatový zápis Dublin Core je možné vytvořit pomocí dvou rozšířených formátů.

První možnost je zapsání záznamu do samostatného souboru ve formátu značkovacího jazyka XML (Extensible Markup Language, rozšiřitelný značkovací jazyk). Jazyku XML je věnován oddíl 5.1.1. Datové služby této práce. Ke každému popisovanému digitálnímu objektu pak existuje jeden soubor metadat. Tato možnost je využívána například pro popis dat v archívech. V praxi bývá metadatový zápis uložen v databázi příslušného archívu a pro potřeby sdílení je vygenerován XML soubor. (Taheri, 2012)

Druhá možnost je zapsání metadat přímo do popisovaného souboru. Popisovaným souborem bývá především webová stránka v HTML. (Soundararajan, 2010) Metadata se pak zapisují do sekce „head” pomocí HTML značky „<meta>”, například:

```
<meta name="DC.Creator" content="Stočes, M.">
```

Zdrojové kódy v následujících kapitolách budou uvedeny pomocí zápisu ve formátu XML. Jako ukázkový záznam bude použita kniha *Babička* od Boženy Němcové.

Soubor patnácti metadatových prvků Dublin Core lze rozdělit do tří logicky členěných skupin: *obsah zdroje, intelektuální vlastnictví a identifikační údaje zdroje*. (Celebová, 2013) (Alilijani, 2009)

Obsah zdroje

Skupina prvků DC je využívána k popisu obsahu zdroje. Jedná se o následující prvky:

Název (dc:title): jméno daného zdroje. Kvalifikátor *Alternative*.

```
<dc:title>
    Babička
<dc:title>
<dcterms:alternative>
    Obrazy z venkovského života
</dcterms:alternative>
```

Předmět, klíčová slova (dc:subject): téma obsahu zdroje. Obsahem prvku bývají většinou klíčová slova. Je doporučeno používat klíčová slova obsažená ve vybraném řízeném slovníku. V následujícím příkladu je využito řízeného slovníku z oblasti zemědělství AGROVOC. Více se řízenými slovníky zabývá oddíl 4.3 Řízené slovníky deskriptorů této práce.

```
<dc:subject xsi:type="dcterms:AGROVOC">
    venkovská populace
</dc:subject>
```

Popis, anotace (dc:description): vysvětlení obsahu zdroje. Zpravidla bývá řetězec volného textu. Používané kvalifikátory jsou: *abstract* (abstrakt), *tableOfContent* (obsah).

```
<dcterms:abstract>
    Babička přijíždí ke své dceři a pomáhá vychovávat vnoučata,
    stává se rádkyní sousedům a nakonec i paní kněžně...
<dcterms:abstract>
```

Typ zdroje (dc:type): typ, kategorie zdroje. Hodnota typu zdroje by měla být vybrána z řízeného slovníku, např. ze slovníku DCMI typů (Dublin Core Metadata Initiative) nebo z PHMK (předmětová hesla národní knihovny). Více se touto problematikou zabývá oddíl 4.3 Řízené slovníky deskriptorů této práce.

```
<dc:type xsi:type="dcterms:DCMIType">text</dc:type>
```

Vztah (dc:relation): odkaz na příbuzný zdroj. Jsou používány kvalifikátory k vyjádření vztah k jinému zdroji. Mezi kvalifikátory patří např. *IsPartOf* (je částí), *IsVersionOf* (je verze), *isFormatOf* (je formátem).

```
<dcterms:isFormatOf>  
    Česká literatura, romantismus.  
</dcterms:isVersionOf>
```

Pokrytí (dc:coverage): rozsah nebo záběr obsahu zdroje. Jedná se především o prostorové nebo časové umístění zdroje.

```
<dc:coverage>19. století</dc:coverage>
```

Intelektuální vlastnictví

Metadatové prvky slouží k identifikaci vlastnického vztahu k jiným objektům.

Prvky Dublin Core:

Tvůrce (dc:creator) obsahuje entitu primárně odpovědnou za vytvoření obsahu zdroje. Pro zvýšení vyhledatelnosti dokumentu je doporučeno používat následující syntaxi pro zápis jména osoby: "Příjmení, Křestní jméno".

```
<dc:creator>Božena Němcová</dc:creator>
```

Vydavatel (dc:publisher) obsahuje entitu odpovědnou za zpřístupnění zdroje.

```
<dc:publisher>Jaroslav Pospíšil </dc:publisher>
```

Příspěvatel (dc:contributor) obsahuje entitu, která přispěla k vytvoření obsahu zdroje a není jejím tvůrcem. Např. se jedná o editory, překladatele aj.

```
<dc:contributor>František Jan Mošner </dc:contributor>
```

Práva, správa autorských práv (ds:rights) informace o právech a licencích, vztahujících se k popisovanému zdroji. Kvalifikátory *rights* (práva), *licence* (licence), *accessRights* (přístupová práva).

```
<dcterms:rights> All rights reserved...</dcterms:rights>
```

Identifikační údaje zdroje, instance zdroje

Jedná se o prvky:

Datum (dc:date) Datum spojené s určitou událostí během existence zdroje. Mezi kvalifikátory prvku datum patří například *created* (datum vytvoření), *dateSubmitted* (datum vložení, podání), *modified* (datum úpravy). Datum se zadává především ve formátu W3C DTF (*Date and Times Format*), vycházející ze standardu ISO 8601:1988.

```
<dcterms:created xsi:type="dcterms:W3CDTF">
    05-01-1855
</dcterms:created>
```

Formát (dc:format) vyjadřuje fyzická nebo digitální reprezentace zdroje. Je doporučeno zapisovat formát digitálních objektů pomocí standardu IMT (Internet Media Type) (W3C, 2002). Jako obsah prvku formát se používají kvalifikátory *extent* (přípona, typ souboru), *medium* (médium – CD, DVD, ...).

```
<dc:format xsi:type="dcterms:IMT"> text/plain </dc:format>
```

Identifikátor zdroje (dc:identifier) obsahuje jednoznačný odkaz na zdroj v rámci daného kontextu. Jedná se například o ISBN (mezinárodní standard číslování knih), ISSN (mezinárodní standard číslování periodické publikace).

```
<dc:identifier xsi:type="dcterms:URI">
    http://search.mlp.cz/cz/titul/babicka/3347644/
</dc:identifier>
```


Jazyk (dc:language) Jazyk intelektuálního obsahu zdroje. Kód jazyku zapisuje především dle normy ISO 639 verze 2, 3.

```
<dc:language xsi:type="dcterms:ISO639-3">ces</dc:language>
```

4.2.1.2 Slovník DCMI Typů

V kvalifikovaném zápisu Dublin Core se využívají pojmy DCMI (Dublin Core Metadata Initiative). Mezi kvalifikátory typů DCMI patří *Collection* (sbírka), *Dataset* (databáze), *Event* (událost), *Image* (obrázek), *InteractiveResource* (multimediální, interaktivní), *MovingImage* (pohyblivý obrázek, video), *PhysicalObject* (fyzický objekt), *Service* (služba), *Software*, *Sound* (zvuk), *StillImage* (statický obraz) a *Text*. (DCMI, 2010)

4.2.1.3 Formáty MARC

MARC (MACHINE-Readable Cataloging) představuje standardy, sestávající z MARC formátů (viz tabulka 3), které slouží k strojově čitelné katalogizaci, ukázka kódu viz obrázek 4. Formáty byly vytvořeny v 60. letech v Kongresové knihovně USA. Struktura záznamu MARC je implementací standardu ISO 2709, známého také jako ANSI/NISO Z39.2. Datový obsah záznamů je definován jinými standardy, jako například AACR2, LCSH nebo MeSH. MARC řeší komplexně problematiku strojově čitelné katalogizace, ale pomocí již zastaralých technologií z tohoto důvodu není již v současné době příliš používán. Reakcí na zastaralost technologií je v současné době využíván formát MARC 21 pro efektivnější výměnu katalogizačních informací. (MARBI, 1996) (Taylor, 2009)

MARC formát	Popis
Autorizační	Poskytuje informace o jednotlivých jménech, předmětech a jednotlivých titulech.
Bibliografický	Popisuje myšlenkové a fyzikální vlastnosti bibliografických zdrojů (knih, zvukových záznamů, zvukových nahrávek apod.).
Klasifikační	MARC záznamy obsahující klasifikaci dat.
Komunitní informace	MARC záznamy popisující zdroj poskytovaných služeb.
Vlastnický	Poskytovat informace o výtisku (katalogizační číslo, umístění police, počet kusů, atd.)

Tabulka 3 MARC – formáty. Zdroj: (Taylor, 2009).

```
=LDR 05594cam 2200565 i 4500
=001 ocn798437851
=003 OCoLC
=005 20131018101403.0
=006 c|||||d|||||
=007 cr|||||
=008 120705s2013||||mauaf\\|b\\|001\0\eng\\
=020 \\$z9780262018555 (hardcover : alk. paper)
=020 \\$z0262018551 (hardcover : alk. paper)
=040 \\Serda
=050 14$BF503$b.A28 2013
=082 04$a153$223
=245 00$aAction science :$bfoundations of an emerging discipline /$cedited by Wolfgang Prinz, Miriam Beisert, and Arv.
=264 \1$aCambridge, Mass. :$bMIT Press,$c[2013]
=264 \4$a@2013
=300 \\$a1 online resource (xi, 450 pages, 5 unnumbered pages of plates) :$billustrations (some color)
=336 \\$atext$btxt$2rdacontent
=337 \\$aunmediated$bn$2rdamedia
=338 \\$aonline resource$bcr$2rdacarrier
=504 \\$aIncludes bibliographical references and index.
=505 0\\$aAcknowledgments -- Contributors -- Action science emerging : introduction and leitmotifs / Arvid Herwig, Mir.
=520 \\$aThe emerging field of action science is characterized by a diversity of theoretical and methodological appro.
=588 \\$aDescription based on print version record.
=650 \0$aMotivation (Psychology)
=650 \0$aCognitive psychology.
=700 1\\$aPrinz, Wolfgang,$d1942-
=700 1\\$aBeisert, Miriam,$d1980-
=700 1\\$aHerwig, Arvid,$d1979-
=776 08$iPrint version: $tAction science.$dCambridge, Mass. : MIT Press, [2013]$w(DLC) 2012024586$w(OCoLC)798437851
=856 40$uhttp://cognet.mit.edu/book-detail/9780262312974$zMITCogNet
```

Obrázek 5 – Ukázka záznamu ve formátu MARC (Abrahamse, 2013)

4.2.1.4 Schéma MODS

Metadata Object Description Schema (MODS) (metadatové schéma pro popis) byl vyvinut odborníky pod hlavičkou organizací knihovna kongresu USA a Marc Standard Office jako podmnožina MARC. První verze metadatového schématu MODS vyšla v roce 2002. Schéma slouží digitálním

knihovnám k popisu jakéhokoliv dokumentu pomocí souboru XML. Schéma MODS je tvořeno 20 prvky (viz tabulka 4), které jsou přejaty z jiných metadatových schémat, díky tomu je schéma konvertovatelné do MARC 21 a Dublin Core. (MODS, 2009) (Švástová, 2009)

<ul style="list-style-type: none"> • titleInfo (nadpis) • name (jméno) • typeOfResource (typ zdroje) • genre (žánr) • originInfo (původní zdroj) • language (jazyk) • physicalDescription (fyzický popis) • abstract (abstrakt) • tableOfContents (obsah) • targetAudience (cílová skupina) • note (poznámka) 	<ul style="list-style-type: none"> • subject (předmět) • classification (klasifikace) • relatedItem (související položky) • identifier (identifikátor) • location (lokace) • accessCondition (přístupová práva) • part (část) • extension (rozšíření) • recordInfo (informace o záznamu)
--	---

Tabulka 4 – Hlavní prvky schématu MODS. Zdroj: (MODS, 2009).

4.2.1.5 Datový model LOM

Learning Object Metadata (LOM) (metadatový model učebních objektů) je standardizovaný model vytvořený pro popis objektů sloužících k učení. Model je definován otevřeným standardem The IEEE 1484.12.1 – 2002 vytvořeným organizací IEEE (Institute of Electrical and Electronics Engineers Standards Association, New York). Relevantní atributy „učebních objektů“, které mají být popsány, zahrnují: *typ objektu; autora; vlastníka; podmínky distribuce; formát*; A pedagogické atributy, jako je styl učení nebo interakce.

Learning Object Metadata je datový model, obvykle kódovaný v XML, používaný k popisu učebního objektu a podobných digitálních zdrojů používaných k podpoře učení. Účelem metadat učení objektu je podporovat opakovanou použitelnost učebních objektů, usnadňovat jejich objevitelnost a usnadňovat jejich interoperabilitu, obvykle v kontextu online systémů řízení učení (LMS).

Další metadatové formáty

Mezi další metadatové formáty patří například Darwin Core, EBU Core, PBCore, CDWA-Lite/CCO, EAD, DACS, ISAD(G), VRA Core, SPECTRUM (Collective Access, 2014).

V oblasti katalogizace vědeckých prací a digitálních knihoven je v současné době nejrozšířenější a nejpoužívanější metadatový formát DC.

4.2.2 Aplikační profil

Při vytváření znalostní databáze, která bude obsahovat metadata objektů je třeba definovat její strukturu. Definováním metadatových elementů, pravidel pro hodnoty a řízených slovníků vzniká tzv. *Aplikační profil* (AP). Velká skupina *Aplikačních profilů*, sloužících k popisu objektů využívá prvky z Dublin Core a rozšiřuje je na základě požadavků aplikace. Příkladem může být například *VOA3R aplikační profil* (Virtual Open Access Agriculture and Aquaculture Repository Application Profile) (Sgouropoulou, 2011) nebo EVSKP-MS (Metadatový soubor pro elektronické vysokoškolské kvalifikační práce v ČR). (Bratková, 2008) „AP je schéma metadat, která se skládá z prvků metadat vybraných z jedné nebo více standardních schémat metadat a je vytvořena tak, aby umožňovala dané aplikaci splnit její funkční požadavky.“ (HEERY, 2008) Evropský výbor pro normalizaci (European Committee for Standardization) (CEN) definuje AP jako „Soustava prvků metadat vybraných z jedné nebo více schémat metadat a kombinovaná ve složeném schématu. Profily aplikací poskytují prostředky k vyjádření

principů modularity a rozšiřitelnosti. Účelem aplikačního profilu je přizpůsobit nebo kombinovat existující schémata do balíčku, který je přizpůsoben funkčním požadavkům konkrétní aplikace, přičemž zachovává interoperabilitu s původními základními schématy.“ (SMITH, 2016)

Dle IMS GLC (2017) jsou důvody tvorby nových aplikačních profilů následující:

- Splnit technické a další požadavky a preference specifické pro projekt, komunitu, doménu nebo region.
- Řešení nejednoznačnosti a obecnosti ve specifikaci nebo standardu.
- Podporovat sémantickou interoperabilitu, např. Pomocí běžně známých slovníků.
- Pro usnadnění testování shody a úspěšné spolupráce.

4.2.2.1 Postupy tvorby aplikačních profilů

V odborné literatuře jsou zmíněny dva významné postupy pro vytváření aplikačních profilů (Guidelines for Developing Application Profiles).

Mezinárodní organizace jako IMS GLC a Evropský výbor pro normalizaci (CEN) (Comité Européen de Normalisation)) vydaly pokyny pro vývoj AP se zvláštním zaměřením na standard IEEE LOM. Tyto pokyny zahrnují následující kroky:

- výběr datových elementů,
- určení počtu výskytu elementů,
- definice datového typu prvků,
- přidání lokálních datových prvků,
- určení povinnosti elementů,
- definování rozsahu hodnot,
- určení závislostí a vztahů mezi elementy,
- určení profilů datových typů,

- vytvoření aplikačního profilu. (Smith, 2016)

„Singapurský Framework“ (SF) (Singapore Framework) je postup pro návrh aplikačních profilů pro Dublin Core. Aplikační profil se dle SF skládá z komponent: (Nonthakarn, 2015)

Funkční požadavky

Funkční požadavky Dublin Core jsou navrhovány k popisu funkcí aplikačního profilu a jejich podpoře. Funkční požadavky tvoří základ pro hodnocení aplikačního profilu pro vnitřní konzistenci a pro poskytování návodů na vhodnost aplikačního profilu pro dané použití.

Model domény

Doménový model definuje základní elementy popsané v profilu aplikace a jejich základních vztazích. Účelem modelu domény je definovat základní oblast aplikačního profilu. Doménový model lze vyjádřit pouze pomocí textu nebo pomocí formálnějšího přístupu, jako je UML.

Sada profilu

Popis sady profilu (Description Set Profile) (DSP) definuje sadu záznamů metadat, které jsou platnými instancemi profilu aplikace.

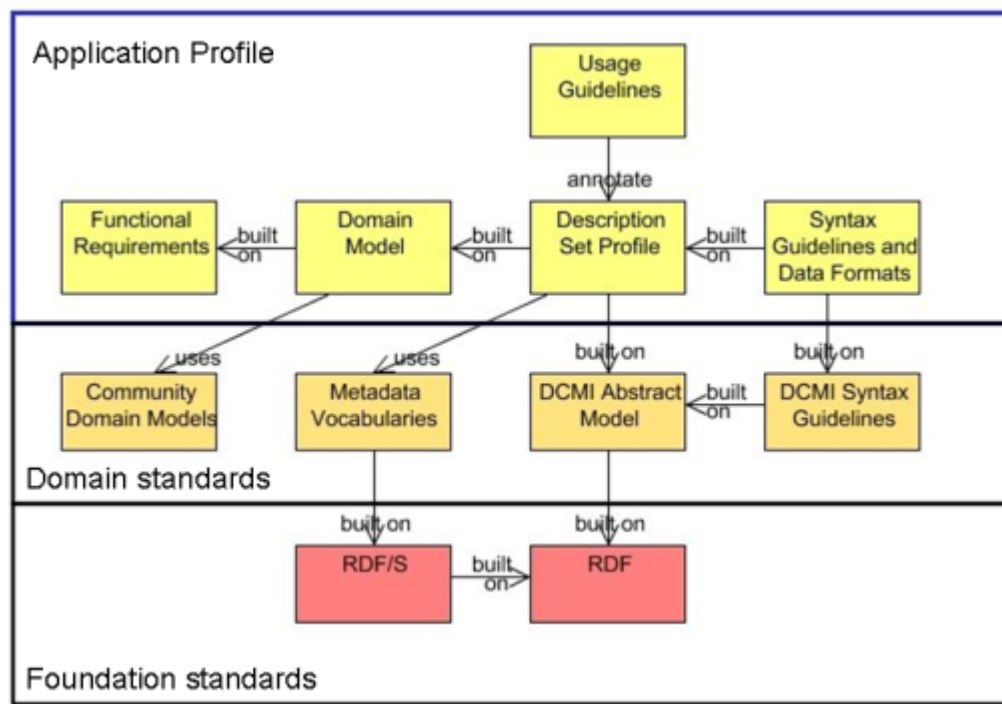
Pokyny pro použití

Volitelné pokyny k použití popisují, jak aplikovat profil aplikace, jak mají být použité vlastnosti v kontextu aplikace apod.

Pokyny pro kódování syntaxe

Volitelné pokyny pro kódování syntaxe popisují všechny syntaxe specifických pro aplikační profil a / nebo pokyny pro syntaxi, pokud existují. (Nilsson, 2009)

Tento model je znázorněn na následujícím obrázku 6.



Obrázek 6 – Součásti profilu Dublin Core Application (Singapore Framework, 2017)

VOA3R AP

Příkladem aplikačního profilu je např. formát VOA3R Metadata AP, který byl vyvinut speciálně pro popis a sdílení dat vědeckých a výzkumných zdrojů z oblasti zemědělství, vodohospodářství, životního prostředí a rozvoje venkova v rámci projektu Virtual Open Access Agriculture and Aquaculture Repository. (Sgouropoulou, 2011)

Pro komplexní popis dat nebo objektu postačuje udržovat záznamy v povinných a velmi doporučovaných elementech. Pro podrobnější charakteristiku je vhodné zaznamenat i doporučované, popřípadě volitelné elementy (Šimek, 2013). Tento metadatový formát částečně vychází ze standardu DC (Šimek, 2012). VOA3R AP 1 se shoduje s DC. K popisu použitých metod, technik a klíčových slov využívá VOA3R AP AGROVOC. Seznam prvků VOA3R AP se nachází v tabulce 2.

Povinné	Silně doporučené	Doporučené	Volitelné
title	creator	description	alternativeTitle
date	contributor	bibliographicCitation	abstract
language	publisher	accessRights	relation
type	identifier	licence	conformsTo
name	format	rights	references
	isShownBy	reviewStatus	isReferencedBy
	isShownAt	publicationStatus	hasPart
	subject	hasMetametadata	isPartOf
	firstName	personalMailbox	hasVersion
	lastName	objectOfInterest	isVersionOf
		variable	hasTranslation
		method	isTranslationOf
		protocol	
		instrument	
		techniques	

Tabulka 5 - Přehled elementů VOA3R AP 4 zdroj (Sgouropoulou, 2011).

4.2.3 Řízené slovníky deskriptorů

Řízené slovníky deskriptorů bývají též označovány jako *Tezaurus*. Tezaurus je referenční příručka, druh slovníku, který uživateli nabízí seznam synonym, někdy i antonym, často jsou mezi pojmy slovníku definované závislosti.

„Řízený slovník deskriptorů, mezi nimiž jsou určeny vztahy nadřazenosti a podřazenosti, termíny synonymní a jiné související. V odborné literatuře popsán jako řízený a měnitelný slovník deskriptorového a selekčního jazyka uspořádaný tak, že explicitně zachycuje apriorní vztahy mezi lexikálními jednotkami.“ (Easylibrary, 2010)

Existuje celá řada oborově a účelově zaměřených řízených slovníků.

4.2.3.1 Rogetův tezaurus

„Rogetův tezaurus (Roget's Thesaurus) je široce užívaný tezaurus anglického jazyka, vytvořený poprvé Peterem Markem Rogetem v roce 1805 a vydaný roku 1852. Původní vydání má 15 000 slov – každé nové vydání je o něco rozsáhlejší. Původní rukopis má ve svých sbírkách knihovna Karpeles Manuscript Library (Karpelesova knihovna rukopisů).

Rogetův Tezaurus se skládá z šesti prvotních tříd (class). Každá třída se skládá z mnoha oddílů (division) a ty zase z částí (section). To je možno chápat jako strom, zahrnující více než tisíc větví (branch) pro jednotlivé „významové skupiny“ (meaning clusters), tedy skupiny slov dle jejich významu či tématu. Jeden z nejobecnějších pojmů významové skupiny je znázorněn jako klíčové slovo v záhlaví, pod nímž jsou ve slovníku řazena všechna slova skupiny.“ (Lloys, 1982), (Hüllen, 2004)

4.2.3.2 Tezaurus AGROVOC

AGROVOC je rozsáhlý tezaurus, vyvinutý výzkumnými pracovníky Organizace pro výživu a zemědělství (FAO), která je specializovanou agenturou Organizace spojených národů (OSN) v rámci AIMS (Agricultural Information Management Standards). AGROVOC obsahuje slova z oblastí potravin, výživy, zemědělství, rybolovu, lesnictví a životního prostředí. Tezaurus obsahuje více než 32 tisíc pojmů v 23 jazycích (duben 2017) včetně českého jazyka. Celý tento tezaurus je vyjádřen jako koncepční systém SKOS (Simple Knowledge Organization System) a publikován jako Linked Data (propojená data), což představuje datový model pro reprezentaci strukturovaných slovníků. Koncepční schéma tezauru AGROVOC používá tři úrovně zobrazení:

- pojmy mají abstrakční význam a jsou zároveň často popsány pomocí adresy URI (Uniform Resource Identifier), např. pro hovězí maso je použito:

http://aims.fao.org/aos/agrovoc/c_861

- pojmy jsou specifikovány jazykově, např.:
البقر لحوم (Arabsky), 牛肉 (Čínsky), hovězí maso (Česky), Viande bovine (Francouzky).
- pojmy mají specifické možnosti (rozsahy) jako jsou pravopisné varianty nebo jednotná a množná čísla, např.:
slepice, slepice, kráva, krávy, apod.

Tímto systémem jsou zabezpečeny terminologické vztahy mezi pojmy a konkrétním významem. AGROVOC je tak vhodný pro popis např. vědeckovýzkumných článků, odborných článků, informací či zpráv z oblasti agrárního sektoru, audiovizuálních dat, apod. (Šimek, 2013)

Českou verzi zpracoval Ústav zemědělských a potravinářských informací v letech 1995 a 1996 v rámci projektu Vypracování české verze tezauru AGROVOC jako předpoklad tvorby národního zemědělského informačního systému, který byl zadán Ministerstvem zemědělství České republiky. Od roku 1997 se používá při zpracování záznamů do článků bibliografie Česká zemědělská a potravinářská bibliografie. (Zemědělská a potravinářská knihovna, 2013) (Šimek 2013)

products > animal products > meat > beef

PREFERRED TERM	beef																																												
DEFINITION	Meat from an adult domestic bovine																																												
BROADER CONCEPT	meat																																												
RELATED CONCEPTS	baby beef veal																																												
IN OTHER LANGUAGES	<table><tr><td>لحم البقر</td><td>Arabic</td></tr><tr><td>牛肉</td><td>Chinese</td></tr><tr><td>hovězí maso</td><td>Czech</td></tr><tr><td>Viande bovine</td><td>French</td></tr><tr><td><i>Viande de boeuf</i></td><td></td></tr><tr><td>Rindfleisch</td><td>German</td></tr><tr><td>गोमांस</td><td>Hindi</td></tr><tr><td>marhahús</td><td>Hungarian</td></tr><tr><td>Carne bovina</td><td>Italian</td></tr><tr><td>牛肉</td><td>Japanese</td></tr><tr><td>쇠고기</td><td>Korean</td></tr><tr><td>ຊີ້ນງົວ</td><td>Lao</td></tr><tr><td>Daging lembu</td><td>Malay</td></tr><tr><td>گوشت گاو</td><td>Persian</td></tr><tr><td>Wotowina</td><td>Polish</td></tr><tr><td>Carne de bovino</td><td>Portuguese</td></tr><tr><td>говядина</td><td>Russian</td></tr><tr><td>hovadzie mäso</td><td>Slovak</td></tr><tr><td>Carne de res</td><td>Spanish</td></tr><tr><td>ఊజిటాలిన్ లనాటా</td><td>Telugu</td></tr><tr><td>เนื้อโค</td><td>Thai</td></tr><tr><td>siğir eti</td><td>Turkish</td></tr></table>	لحم البقر	Arabic	牛肉	Chinese	hovězí maso	Czech	Viande bovine	French	<i>Viande de boeuf</i>		Rindfleisch	German	गोमांस	Hindi	marhahús	Hungarian	Carne bovina	Italian	牛肉	Japanese	쇠고기	Korean	ຊີ້ນງົວ	Lao	Daging lembu	Malay	گوشت گاو	Persian	Wotowina	Polish	Carne de bovino	Portuguese	говядина	Russian	hovadzie mäso	Slovak	Carne de res	Spanish	ఊజిటాలిన్ లనాటా	Telugu	เนื้อโค	Thai	siğir eti	Turkish
لحم البقر	Arabic																																												
牛肉	Chinese																																												
hovězí maso	Czech																																												
Viande bovine	French																																												
<i>Viande de boeuf</i>																																													
Rindfleisch	German																																												
गोमांस	Hindi																																												
marhahús	Hungarian																																												
Carne bovina	Italian																																												
牛肉	Japanese																																												
쇠고기	Korean																																												
ຊີ້ນງົວ	Lao																																												
Daging lembu	Malay																																												
گوشت گاو	Persian																																												
Wotowina	Polish																																												
Carne de bovino	Portuguese																																												
говядина	Russian																																												
hovadzie mäso	Slovak																																												
Carne de res	Spanish																																												
ఊజిటాలిన్ లనాటా	Telugu																																												
เนื้อโค	Thai																																												
siğir eti	Turkish																																												
URI	http://aims.fao.org/aos/agrovoc/c_861																																												
Download this concept:	RDF/XML TURTLE																																												

Obrázek 7 Ukázka výpisu záznamů "hovězí maso" z tezauru AGROVOC. Zdroj: FAO.

4.2.3.3 Tezaurus EuroVoc

EuroVoc je vícejazyčný tezaurus zaměřený na oblast práva a legislativy Evropské unie (EU). EuroVoc je k dispozici ve 23 úředních jazycích Evropské unie a v jednom jazyce třetí země (srbštině).

Tezaurus byl vytvořen ve spolupráci Evropského parlamentu, Komise EU a Úřadu pro úřední tisky Evropského společenství (ES) za podpory DG XIII (Directorate General). Za správu EuroVocu na mezinárodní úrovni odpovídá Úřad pro úřední tisky ES.

EuroVoc spravuje Úřad pro publikace. Správa tezauru nyní volí ontologický přístup a v souladu s doporučeními konsorcia World Wide Web

(W3C) a s nejnovějšími trendy v tvorbě tezurů používá sémantické webové technologie.

Tezaurus EuroVoc používají mimo jiné tyto instituce: Evropský parlament, Úřad pro publikace, parlamenty členských států EU a jejich regionů, správní orgány členských států a soukromí uživatelé z členských i nečlenských zemí. (EuroVoc, 2015) (Schwarz, 1999)

Mezi další řízené slovníky patří:

- **Moby Thesaurus** – rozsáhlý slovník obsahující přes 30 000 kořenových pojmů a 2,5 milionů synonym.
- **The Art & Architecture Thesaurus (AAT)** – tezaurus obsahující pojmy z umění, architektury, archeologie a kulturního dědictví.
- **ITIS** (Integrated Taxonomic Information System) – tezaurus taxonomie rostlin a živočichů.
- **MeSH** (Medical Subject Headings) – slovník využívaný v medicíně.
PHMK MESH PSH
- **TGN** (Thesaurus of Geographic Names) – tezaurus geografických názvů.

Většina tezurů obsahuje termíny pouze v anglickém jazyce. Přeložení tezurů je časově a finančně náročné. Většina vědeckých publikací je prezentována v anglickém jazyce a proto není potřeba slovníky překládat.

Kromě tezurů existují i mezinárodní předmětové klasifikace.

4.2.3.4 Předmětová klasifikace Jamese Duffa Browna

Předmětová klasifikace vytvořená v roce 1894 Jamesem Duffem Brownem a stala se počátkem moderních bibliografických klasifikačních systémů. Brown si jako jeden z prvních uvědomoval nutnost interdisciplinárního přístupu a zastával názor, že nejlepší metodou pro

vyhledávání materiálů v knihovně je klasifikovaný katalog. Základní obecné kategorie klasifikace jsou:

- A Religion and philosophy (náboženství a filozofie)
- B History, travel and topography (historie, cestování a topografie)
- C Biography (bibliografie)
- D Social science (sociální vědy)
- E Science (věda)
- F Fine and recreative arts (krásné umění)
- G Useful arts (užité umění)
- H Language and literature (jazyk a literatura)
- J Poetry and the drama (poezie a drama)
- K Fiction (fikce)
- L General works (všeobecné)

4.2.3.5 JEL klasifikace

Klasifikační systém byl vyvinut pro použití v Journal of Economic Literature (JEL) a je standardní metodou klasifikace odborné literatury v oblasti ekonomie. Systém se používá ke klasifikaci článků, disertací, knih, recenzí knih a pracovních dokumentů v EconLit a v mnoha dalších aplikacích.

4.2.4 Webové služby

Webová služba je metoda komunikace mezi dvěma elektronickými zařízeními prostřednictvím sítě. Specifikace webové služby podle konsorcia W3C (W3C, 2004): „Webová služba je softwarový systém, vytvořený k podpoře interakce mezi stroji probíhající přes počítačovou síť. Má rozhraní popsané ve strojově zpracovatelném formátu (specificky WSDL) (Web Services Description Language). Ostatní systémy interagují s webovou službou způsobem předepsaným jejím popisem prostřednictvím SOAP (Simple Object Access Protocol) zpráv, typicky přenášených pomocí HTTP

(Hypertext Transfer Protocol) s XML serializací v součinnosti s ostatními webovými standardy.“ (Kuba, 2006)

Koncept webových služeb vznikl jako reakce na potřebu obecnějšího řešení při komunikaci v distribuovaných systémech, tedy systémech, jejichž části běží na různých počítačích, spojených komunikační sítí. (Møller, 2006) Hlavním principem webových služeb je umožnit programátorům vzdáleně volat funkce, fyzicky běžící na jiných strojích prostřednictvím počítačové sítě a to tak, že vzdálené volání funkcí se nebude lišit od lokálního volání. Koncept webových služeb využívá myšlenku vzdáleného volání funkcí, ale zároveň si klade za cíl být co nejobecnější a na platformě nezávislý. (Mlýnková, 2008) (Møller, 2006)

4.2.5 Technologie webových služeb

Dle původní specifikace webové služby se jedná o webové služby v souvislosti s několika technologiemi, které je třeba představit. Specifikace zmiňuje formát XML, protokol SOAP a formát WSDL, ale u webových služeb je často používán i název registr webových služeb UDDI (Universal Description, Discovery and Integration).

4.2.5.1 Datové služby

XML

Extensible Markup Language (XML) je značkovací jazyk vyvinutý W3C (World Wide Web Consortium), který vznikl zjednodušením značkovacího jazyka SGML (Standard Generalized Markup Language) (ISO 8879:1986), oficiální verze byla představena v roce 1998. Textový formát XML definovalo konsorcium W3C pro přenos obecných dokumentů a dat. Vznik jazyka XML byl inspirován potřebou univerzálního otevřeného standardu, který by byl schopen jednoduchým způsobem popsat obecnou strukturu dat. Důležitou vlastností XML je jeho přenositelnost mezi různými platformami a národními

prostředími, a také nezávislost na konkrétním programovacím jazyku. (Mlýnková, 2008) V současné době je verze jazyka XML 1.1. (W3C, 2016)

Zápis jazyka je realizován do souborů pomocí prostého textu. Zápis není binární, není závislý na platformě, kde vznikl a je čitelný pro uživatele i stroje. Mezi nejrozšířenější využití jazyka XML patří:

- Prostředek pro výměnu dat mezi systémy.
- Nástroj pro univerzální publikaci dokumentů, lze snadno transformovat do jiného formátu.
- Soubory XML se používají jako konfigurační soubory aplikačních programů.
- Z XML jazyka jsou odvozeny další formáty, např. formát pro vektorovou grafiku SVG (Scalable Vector Graphics) nebo velká množina GIS (Geografický informační systém) formátů.
- Soubory napsané v jazyce XML se využívají pro vývoj grafického rozhraní aplikací např.: OS (operační systém) Android.

Při ukládání dat do nativního XML formátu vzniká velké množství redundantních dat, způsobené častým opakováním XML značek. Za tuto vlastnost je jazyk kritizován. Další nevýhodou tohoto jazyka je jeho strojové zpracování (parsování), které může být u rozsáhlého dokumentu poměrně náročné. (Mlýnková, 2008)

Mezi konkurenty značkovacího jazyka XML patří další formáty pro serializaci dat, např. JSON (*JavaScript Object Notation*), YAML, S-Expressions.

JSON

JSON (JavaScript Object Notation) je datový formát, používaný pro serializaci dat. Jeho specifikace je definována doporučením RFC 7159. Struktura jazyka je jednoduchá a člověkem i strojem čitelná. Struktura JSON je založena na transformaci vstupních datových hodnot (číslo, řetězec, logická hodnota, objekt nebo z nich složené pole) do textového řetězce,

formovaném dvojicemi atribut – hodnota. (Sriparasa, 2013). Využívání datového formátu JSON neustále roste, je to způsobené rozvojem mobilních aplikací, kde je tento formát využíván pro přenos dat mezi serverem a aplikací.

XML a JSON ukládají data do stromové struktury.

CSV

CSV (Comma-separated values) je jednoduchý textový formát pro uložení tabulkových dat. CSV je definováno dle doporučení RFC (Request for Comments). (RFC 4180, 2005)

Příklady zdrojových kódů XML, JSON a CSV

```
<kniha>
  <nazev>Babička</nazev>
  <autor>Božena Němcová</autor>
  <rok_vydani>1855</rok_vydani>
</kniha>
```

(XML)

```
{
  "kniha": {
    "nazev": "Babička",
    "autor": "Božena Němcová",
    "rok_vydani": "1855"
  }
}
```

(JSON)

```
Babička, Božena Němcová, 1855
```

(CSV)

Popis formátů

Pro popis formátů se využívají tzv. schémata. Pro schémata dokumentů ve formátu XML se používají dvě schémata DTD (document type definition) a pokročilejší schéma XML, označováno jako XSD (XML Schema Definition) (de Almeida, 2016). XSD bylo publikováno jako doporučení konsorciem W3C. Schéma pro datový formát JSON nebylo doposud publikované ve finální verzi. (Barbaglia, 2017)

RDF

RDF je zkratkou pro Resource Description Framework (systém popisu zdrojů). Jak již bylo zmíněno, jedná se o soubor specifikací, které vypracovalo konsorcium W3C. RDF byl původně navržen jako model metadat a v současnosti slouží jako obecná metoda pro modelování informací v různých syntaxích. (W3C, 2017)

RDF přiřazuje jednotlivým zdrojům výrazy ve tvaru podmět (subjekt) – přísudek (vlastnost, predikát) – předmět (objekt). Tyto výrazy se v terminologii RDF nazývají trojice. Podmět určuje zdroj, vlastnost určuje charakter zdroje a vyjadřuje vztah mezi podmětem a předmětem. Tedy například ve výroku: „ČZU je zkratka České zemědělské univerzity“ by byl dle RDF podmět: „ČZU“, vlastnost: „je zkratka“ a předmět: „České zemědělské univerzity“.

Význam této trojici v RDF dávají tzv. slovníky. Tyto slovníky nejsou centralizované a pokud pro určitý obsah slovník neexistuje, může si uživatel napsat slovník vlastní. Je však nezbytné dodržet základní pravidla pro psaní slovníků. Slovníky se vkládají do kódu pomocí XML namespace

```
xmlns:nazev_slovniku="http://adresa_slovniku/
```

a jsou dostupné pouze v XHTML. Jediná schválená specifikace je tak pro XHTML 1.1.

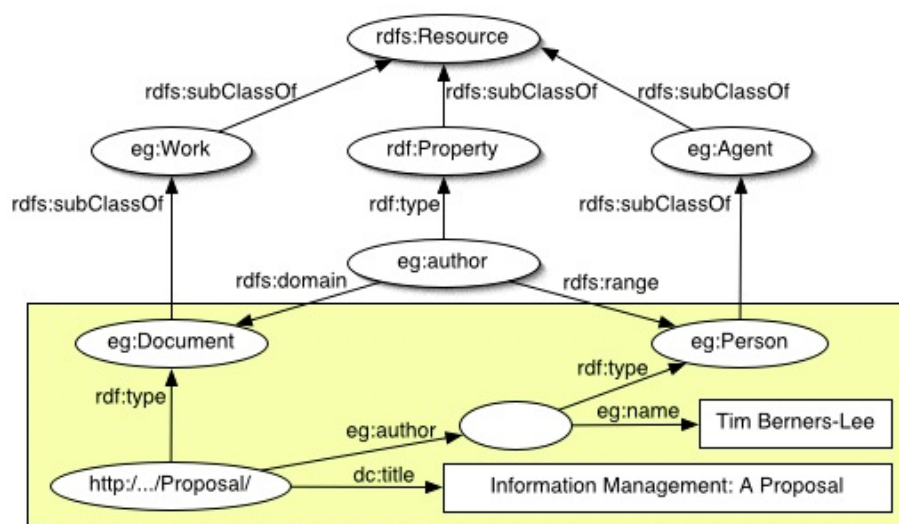
Příklad zdrojový kód RDF

```
<div xmlns:foaf="http://xmlns.com/foaf/0.1/">
<ul>
<li typeof="foaf:Person">
<a property="foaf:name" rel="foaf:mbox"
href="mailto:"e-mail_osoby">jmeno_osoby</a>
</li>
</ul>
</div>
```

Slovník FOAF (friend of a friend) je určen pro popis osob a vztahu mezi nimi. V příkladu je určena osoba a e-mailový kontakt na ni. Stejně tak je možné propojit s konkrétní osobou její webovou stránku, pohlaví, věk aj.

Za pomoci kombinace různých slovníků lze sémanticky popsat široké spektrum jevů. Tato skutečnost přináší RDF velkou výhodu oproti

mikroformátům, které jsou ve svém použití více omezeny. Vzhledem k jejich jednoduchosti a návaznosti na HTML a XHTML jsou však v současné době mikroformáty stále rozšířenější. Z tohoto důvodu není ani podpora RDF ze strany internetových prohlížečů tak dobrá jako u mikroformátů. V současnosti RDF podporuje jen pár zkušebních prohlížečů a jeden doplněk pro prohlížeče Firefox. Podobně je tomu i v oblasti internetových vyhledavačů. RDF dokáže zpracovat jen vyhledavač Google, a to jen pro omezený okruh dat. (Kennedy, 2006)



Obrázek 8. - Ukázka popisu dokumentu RDF (W3C, 2017)

Mikroformáty

Mikrodata jsou nová specifikace v HTML5, která má být použita k zavedení sémantiky v rámci stávajícího obsahu webových stránek. Mikrodata by měla být schopna automaticky extrahovat internetové vyhledavače i prohlížeče. Mikrodata používají podpůrné slovníky pro popis položek a dvojice název–hodnota k přiřazení hodnoty a její vlastnosti.

Mikrodata přidávají do HTML několik nových atributů pro lepší vkládání metadat. Pro identifikaci typů objektů je možné používat URI. Mikrodata také používají vlastní datový model, ale je definován algoritmus pro jeho převod do RDF.

K mikrodatům se váže projekt Schema.org. Tento projekt by měl pomoci poskytovat vyhledavačům lepší výsledky ve vyhledávání a ulehčit jim orientaci v obsahu webových stránek. Klade si za cíl vytvořit návod pro webmastery, jak používat mikrodata určená vyhledavačům. Jedná se o seznam mikrodat, která může tvůrce webu implementovat do svého zdrojového kódu a dát tak datům sémantický význam.

```
<div itemscope itemtype="http://schema.org/Event">
  <div itemprop="name">Thing Together 2017</div>
  <span itemprop="description"> Doktorská vědecká konference, která
  má umožnit studentům prezentovat výsledky vlastní práce a dále
  získat zkušenosti, které lze tvůrčím způsobem využít.</span>

  Event date:
  <time itemprop="startDate" datetime="2017-02-06T9:30">6. února
  2017 od 9:00 hodin</time>
</div>
```

Pomocí "itemscope" je definován celý záznam, v tomto případě událost (event). Pomocí jednotlivých "itemprop" jsou potom definovány konkrétní parametry (název, popis a čas události). Pokud by tato událost byla takto popsána v HTML5, byl by například Google schopen automaticky (strojově) vyhodnotit, že událost Thing Together je Doktorská vědecká konference, a že se koná 6. února 2007 od 9:00 hodin.

Mikrodata vznikla jako kompromis mikroformátů a RDF. Přejímají totiž dvě základní myšlenky obou těchto specifikací. Z mikroformátů přebírají myšlenku jednoduchých atributů (tříd), použitelnou v HTML, a z RDF existenci slovníků. V této specifikaci jsou vytvořeny, na rozdíl od mikroformátů, speciální skupiny atributů (itemscope a itemprop). Není tedy již nutné používat běžné html třídy, které mohou jednotliví tvůrci webů využívat libovolně. Výhodou mikrodat oproti RDF je to, že není nutná specifikace XHTML.

Primární výhodou mikrodat je potom skutečnost, že se je rozhodly podporovat dva významné vyhledavače – Google a Bing – což je základní

předpoklad pro rozšíření jakékoliv sémantické specifikace. Protože na implementaci a využití mikrodat nyní pracují velké vyhledávače, je nasnadě, že je zanedlouho budou následovat i tvůrci webových stránek.

Obecná myšlenka všech tří sémantických specifikací (mikroformáty, RDF a mikrodata) je stejná. Jejich cílem je umožnit strojové zpracování dat vyhledávačům, potažmo prohlížečům v duchu koncepce sémantického webu. Podobný je i způsob, jakým toho chtějí dosáhnout. (Kennedy, 2006)

4.2.5.2 Protokol SOAP

SOAP (*Simple Object Access Protocol*) je jedním z protokolů pro výměnu zpráv mezi webovými komponentami a může být použit jak k výměně XML dat, tak ke vzdálenému volání funkcí. SOAP protokol je založený na XML. To znamená, že každá SOAP zpráva je XML dokument se strukturou popsanou protokolem SOAP. Díky nezávislosti na konkrétní platformě či jazyku je často využíván pro výměnu zpráv mezi webovými službami. Za jeho nevýhodu lze označit přílišnou složitost, která se projevuje poměrně dlouhým zápisem každé SOAP zprávy. (Mlýnková, 2008)

4.2.5.3 Protokol RPC

RPC (Remote procedure call) je vzdálené volání procedur, používané již od roku 1976 dle normy RFC 707, technologie dovolující programu vykonat proceduru, která může být uložena na jiném místě, než je umístěn sám volající program. Příkladem budiž výpočet funkce na jiném počítači v síti. (RFC 1813, 1995) (RFC 707, 1976)

4.2.5.4 Rozhraní REST

REST (Representational State Transfer) je architektura rozhraní, navržená pro distribuované prostředí. Pojem REST byl představen v roce 2000 v disertační práci Roye Fieldinga. REST je, narozdíl od známějších XML-RPC či SOAP, orientován datově, nikoli procedurálně. Webové služby definují vzdálené procedury a protokol pro jejich volání, REST určuje, jak

se přistupuje k datům. REST implementuje čtyři základní metody, které jsou známy pod označením CRUD, tedy vytvoření dat (*Create*), získání požadovaných dat (*Retrieve*), změnu (*Update*) a smazání (*Delete*). Tyto metody jsou implementovány pomocí odpovídajících metod HTTP protokolu. (Webber, 2010)

4.2.5.5 Formát WSDL

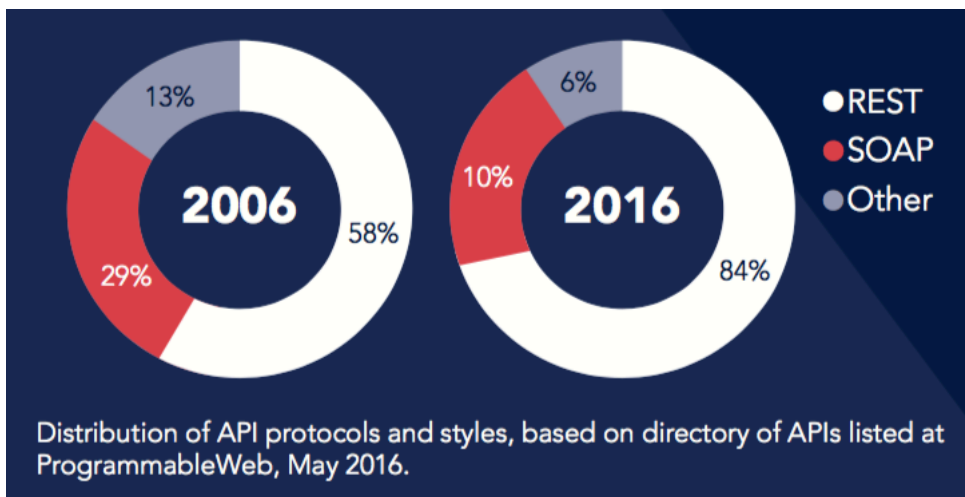
Možnost vzdáleně volat funkce pomocí protokolu SOAP lze využít pouze, pokud je známo rozhraní služby, tj. jaké funkce služba poskytuje a jaké má funkce parametry. WSDL (Web Services Description Language) je jazyk vyvinutý jako standardizovaný popis rozhraní webových služeb a je založen na XML (Mlýnková, 2008). Popis WSDL vymezuje množinu abstraktních operací, určuje, jaké jsou tyto operace povahy (zda jsou obousměrné či pouze jednosměrné), jak tyto operace volat, jaké mají vstupní a výstupní parametry, a jakých jsou datových typů. Současně určuje, kde je webová služba dostupná a jaké mapování (na jaký protokol) se používá pro definované abstraktní zprávy. Popis webové služby WSDL je opět XML dokument s předepsanou strukturou. (Kuba, 2005). V něm se zavádí potřebné jmenné prostory.

Graf 1 autora K. Honga (2015) znázorňuje poměr používaných API (Application Programming Interface) technologií. Z grafu je patrné, že v 3/2014 byla nejrozšířenější technologií REST (69%), dále SOAP (18%), JavaScript (5%) a RPC (2%).



Graf 1 - Poměr API a technologií. Zdroj: (Hong, 2015)

Současným trendem je nárůst Java Scriptových API na úkor tradičních SOAP a RESC (viz Graf 2)



Graf 2 - REST a SOAP api 2006 a2016 Zdroj: <https://www.programmableweb.com>

4.2.6 Technologie pro sdílení dat

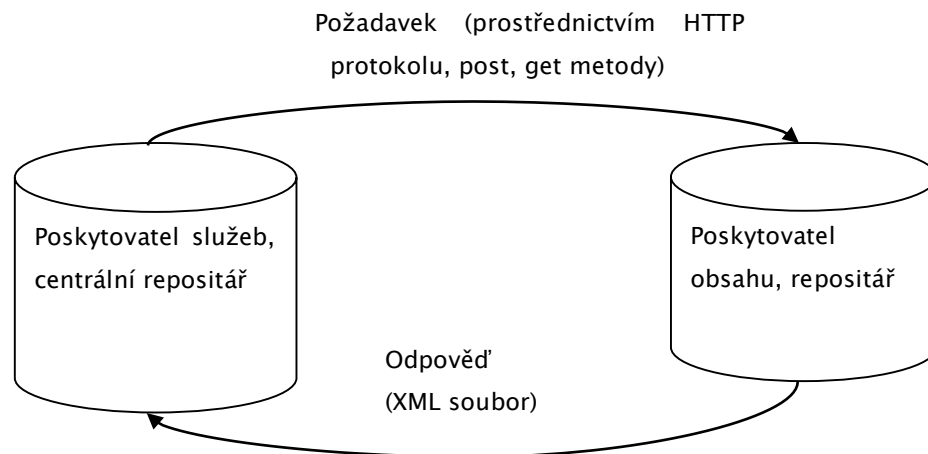
Nejrozšířenějším protokolem pro přenos metadat ve formátu DC mezi systémy je OAI-PMH (Open Archive Initiative, 2013)

4.2.6.1 Aplikační vrstva Z39.50

Z39.50 je síťová komunikační aplikační vrstva, která se používá v knihovnictví jako standard pro bibliografické informační systémy, a je vyvíjena od 70. let. Z39.50 je standardem dle ISO 23950:1998 a ANSI/NISO Z39.50. Aplikační vrstva umožňuje získávat informace z databáze prostřednictvím protokolů TCP / IP (Transmission Control Protocol/Internet Protocol). Z39.50 je ideální pro použití v portálovém prostředí, které umožňuje jednotný přístup simultánního vyhledávání v několika bibliografických informačních systémech. Nástupce Z39.50 je SRU/CQL (Search/Retrieve via URL/ Contextual Query Language), což je standard pro vyhledávání založený na XML.

4.2.6.2 Protokol OAI-PMH

Open Archives Initiative Protocol for Metadata Harvesting umožňuje katalogům a databázím jednoduše sdílet a nabízet svá metadata pomocí protokolů HTTP a XML standardu. Metadata mohou být v různých formátech, ale Dublin Core je nejpoužívanější a doporučený. Aplikace využívající OAI Framework může působit jako poskytovatel dat (metadata provider) nebo poskytovatel služeb (service provider). Poskytovatelé metadat nabízejí svá data o svých datových objektech (digitální knihovna). Poskytovatelé služeb „sklízíjí“ (harvesting), získávají data od poskytovatelů a poskytují služby vyhledávání. (Fegen, 2013) Princip je znázorněn na obrázku 3.

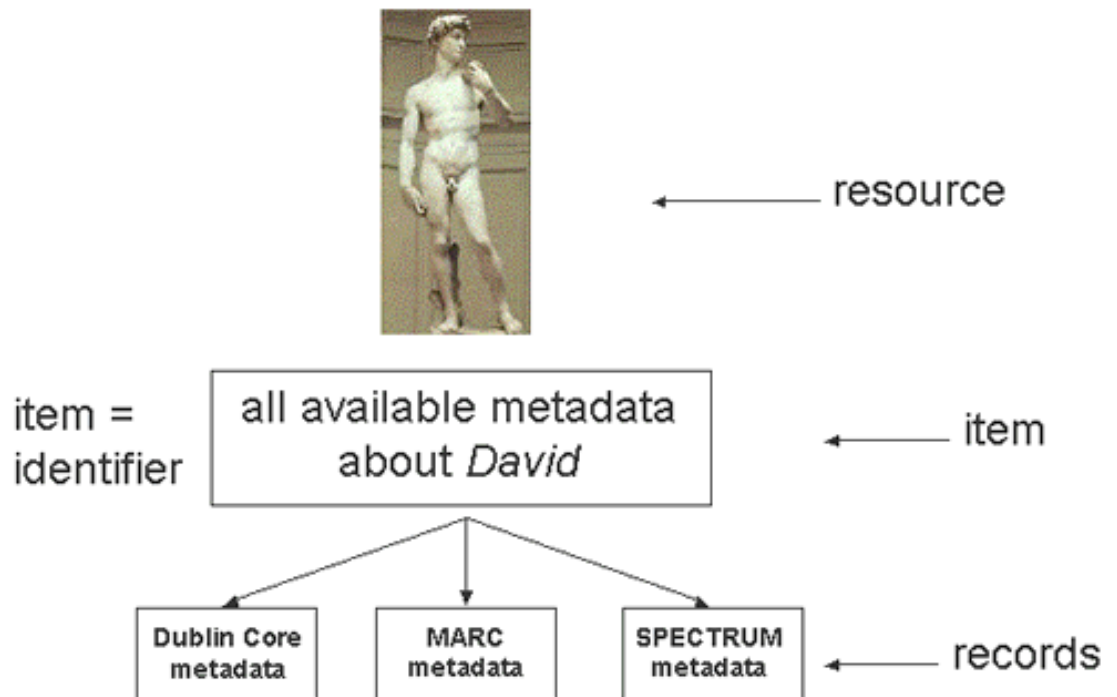


Obrázek 9 – princip OAI-PMH. Zdroj: Autor.

OAI definuje následující pojmy: (Carpenter, 2013)

- **Harvester – sklízeč** – klientská aplikace používající žádosti OAI-PMH.
- **Repository – repositář, úložiště** – síťový server, umožňující odpovídat na žádosti OAI-PMH.
- **Resource – zdroj** – objekt, který je popisován metadaty, může být cokoliv (kniha, digitální článek, socha) – není definován OAI-PMH.
- **Item – položka** – úložiště, ze kterého mohou být metadata o zdroji zpřístupněna, má jedinečný identifikátor.
- **Record – záznam** – metadata v určitém metadatovém formátu.
- **Identifier – identifikátor** – jedinečný klíč pro položku v úložišti.
- **Set – skupina** – volitelná konstrukce pro seskupování položek v úložišti.

OAI pojmy jsou demonstrovány na soše Michelangela Buonarrotiho David na obrázku 4.



Obrázek 10 – OAI-PMH pojmy. Zdroj:(Carpenter, 2013).

4.2.6.3 Druhy OAI-PMH požadavků (request)

Požadavky sklízecí (harvestru) jsou zasílány do repozitáře pomocí HTTP protokolu s využitím metod get a post. Repozitář vrací sklízecí XML dokumenty. V případě zaslání chybného požadavku je vrácené XML s chybovou hláškou výjimkou.

OAI-PMH definuje 6 požadavků, které slouží ke komunikaci mezi repozitářem a harvesterem: (Carpenter, 2013)

- **Identify** – dotaz na identifikaci repozitáře.
- **ListMetadataFormats** – dotaz na dostupné metadatové formáty.
- **ListSets** – dotaz na strukturu seskupení v repozitáři
- **ListIdentifiers** – dotaz na získání záznamů z repozitáře, načte pouze hlavičky.
- **ListRecords** – dotaz na získání záznamů z repozitáře
- **GetRecord** – získat individuální záznam metadat z repozitáře

Princip zasílání informací je znázorněn na obrázku 5.



Obrázek 11 – příklad komunikace mezi centrálním a lokálním repositářem. Zdroj: (Carpenter, 2013)

4.2.6.4 SW řešení OAI-PMH

Existuje celá řada softwarových řešení, která podporují protokol OAI-PMH a umožňují práci s metadaty. Mezi nejznámější patří SW DSpace nebo ePrints (PMH Tools, 2013).

DSpace

DSpace je open source software, sloužící pro vytváření institucionálních repositářů. DSpace umožňuje efektivní správu digitálních a digitalizovaných dat. Podporuje širokou škálu dat, jako například textové dokumenty, 3D digitální objekty, fotografie, videa, výzkumné datové základny apod.

DSpace byl vyvinut v roce 2002 institucemi Massachusetts Institute of Technology (MIT) a HP Labs (laboratoře Hewlett-Packard). V současné době je spravován a vyvíjen neziskovou organizací DuraSpace. DSpace podporuje běžné standardy interoperability, používané v institucionálních repositářích jako je OAI-PMH, SWORD, OpenSearch a RSS. DSpace se ukázal jako nejživější, nejhotovější a nejkomplexnější ze všech systémů

pro provozování digitálních knihoven, které byly na různých univerzitách v ČR testovány. (Donohue, 2015)

ePrints

ePrints vznikl jako reakce na vznik protokolu OAI-PMH. Jedná se o softwarový open-source balík, primárně určený pro budování institucionálních repositářů s podporou OAI-PMH. ePrints je vyvíjen na univerzitě v Southamptonu a je šířen pod licencí GPL. (ePrint, 2016)

The eXtensible Catalog (XC) Drupal Toolkit

The eXtensible Catalog Drupal Toolkit je modulární nástroj pro redakční systém Drupal verze 6, který umožňuje zpracování a manipulaci s metadaty. Obsahuje možnosti pro definování metadat, jejich export, import, ukládání a načítání, pro vyhledávání a úpravy metadat. Dále umožňuje řízení přístupu uživatelů k metadatům, generování plně upravitelného HTML (hypertext markup language) a CSS (Cascading Style Sheets) výstup. Drupal i jeho modul XC Drupal Toolkit jsou open source SW (software) pod licencí GNU (GNU's Not Unix!). (Drupal, 2016)

AgriDrupal

AgriDrupal je rozšíření redakčního systému open-source Drupalu. Doplněk rozšiřuje Drupal o specifické funkce pro správu zemědělských informací. AgriDrupal je vhodný pro malé firmy a instituce jako základní zemědělský informační systém. Rozšíření je součástí platformy AIMS (Agriculture Information Management Standards, Standardy pro zemědělský informační management). (FAO, 2016)

4.2.7 Shrnutí

Pro efektivní práci s digitálními objekty je velmi důležité jejich obohacení o popisná data, tzv. *metadata*.

Při sdílení digitálních dokumentů je sdílen pouze jejich metadatový popis. Velmi rozšířeným a vhodným nástrojem pro sdílení metadat je *OAI-PMH*. OAI-PMH využívá některé principy webových služeb.

Vhodný datový formát pro metadatový zápis je XML. Dále by bylo možné použít JSON, ale transformace je jen technická záležitost. JSON nemá konečnou verzi schémat pro validaci formátu.

Pro další analýzy byly vybrány následující tři metadatové modely vhodné pro popis digitálních artefaktů:

- DC (Dublin Core),
- LOS (Learning object metadata)
- MODS (Metadata Object Description Schema, a bibliographic description schema).

4.3 Možnosti zveřejnění obsahu

Znalosti jsou otevřené, pokud k nim může kdokoliv volně přistupovat, používat je, upravovat a sdílet je – za podmínek, které nejvíce zabezpečí zachování původu, tedy vlastnictví, a otevřenosti. (Open Definition, 2016). Existuje celá řada projektů zabývajících se touto problematikou.

AIMS

AIMS (Agriculture Information Management Standards, Standardy pro zemědělský informační management) jsou standardy, technologie a osvědčené postupy pro open access a open data v oblasti zemědělství. Standardy jsou vytvářené odborníky pod záštitou FAO. Cílem těchto standardů je podpora zemědělských manažerských systémů a odborníků zemědělské správy po celém světě při vytváření přístupu ke znalostem z oblasti zemědělství. (FAO, 2016)

Iniciativa pro otevřenou citaci

Iniciativa pro otevřené citace (Initiative for Open Citation) je projekt, který byl veřejně otevřen v dubnu 2017 a popisuje sebe sama jako: spolupráci mezi vědeckými vydavateli, výzkumnými pracovníky a dalšími zainteresovanými stranami za účelem podpory neomezené dostupnosti vědeckých citačních údajů. (Schiermeier, 2017)

4.3.1 Otevřená data (open data)

Otevřená data jsou data dostupná volně a bezplatně na internetu ve strukturované a strojově čitelné podobě. Formát a struktura otevřených dat by měly umožňovat jejich hromadné zpracování, k němuž jejich vydavatel poskytl právní svolení. Díky tomu mohou být dál volně zpracovávány, a to i v rámci softwarových aplikací (Fond Otakara Motejla, 2016). Publikování otevřených dat je významným trendem nejen v oblasti sémantického webu (Kozaki, 2014). Otevřená data jsou v současné době

podporovanou agendou všech úrovní státní správy, samosprávy a státních nebo veřejných institucí po celém světě. Stoupá poptávka po datech jako takových, což např. dokládá projekt Web informace pro všechny (informaceprovsechny.cz), kde byla většina podaných žádostí žádostmi o poskytnutí dat, nikoli žádostmi o poskytnutí dílčích informací.

Publikování otevřených dat je významný trend nejen sémantického webu, ale projevuje se také v mnoha dalších oblastech, jako jsou společenské vědy, státní správa a samospráva, média, zeměpisné oblasti, publikace, apod. Pro transformaci obsahu a struktury dat jsou rovněž používány různé normy metadat a formáty dat (Spampinato, 2013).

Tim Berners-Lee (2006) rozdělil formát otevřených dat do pěti kategorií, označených příznačně počtem hvězdiček. Nejvyšší počet hvězdiček udává nejpřístupnější formát otevřených dat, jedna hvězdička naopak nejméně strojově přístupná data. Celkový přehled vypadá následovně:

- jedna hvězdička: naskenovaný dokument v pdf (jako obrázek)
- dvě hvězdičky: data v souboru xls (MS Excel)
- tři hvězdičky: data v csv (neproprietární formát)
- čtyři hvězdičky: data v otevřeném formátu W3C
- pět hvězdiček: data jsou navíc v rámci kontextu propojena linky s dalšími otevřenými daty

Výrazným problémem je dnes především velká roztržitost formátů a tím i přístupnost a neexistence jednotného formátu a systému struktury otevřených dat pro jednotlivé oblasti společenského života. Jedním z požadavků na otevřená data je sice jejich strojová čitelnost, reálně jsou však data publikována v heterogenních formátech s výrazně rozdílnou strukturou. Mnoho přístupů je založeno na konkrétním případě a přenos je omezen v důsledku nekompatibilních formátů dat (Klug, 2014). Datové soubory jsou publikovány v proprietární struktuře, což v řadě případů ztěžuje objevit užitečnou sadu dat a vyžaduje uživatelský vklad struktury a sémantiky (Hofman, 2014). V současné době existuje několik pokusů o standardizaci

formátů dat, např. publikování různých sad datových zdrojů v RDF (Resource Description Framework) (Silva, 2013), žádný však problematiku neřeší komplexně, napříč obory a včetně struktury, validace a transformace. Mezi další problémy patří vysoká úroveň redundance a slabá možnost strojově zpracovatelných možností popisu datových souborů (Fernández, 2013). Obsah dat, jejich formát, struktura i kvalita se velmi liší, a to i v případě, kdy jsou data publikována pomocí principů pro široce provázaná data (Böhm, 2012).

Z tohoto důvodu je nutné při analýze a správě rozsáhlých dat vytvářet různé konverze formátů. Důsledkem pak je často náročné řešení v podobě implementace těchto struktur a formátů do aplikací, které jsou využitelné koncovými uživateli. Aby pro ně byla otevřená data ve velké míře efektivně využitelná, je nutné je zobrazovat nebo zpracovávat v uživatelsky komfortních SW aplikacích (desktopových, mobilních). Pak se teprve dostaví výstup v podobě „Méně práce pro úředníka, více pohodlí pro uživatele“.

4.3.2 Licence

Licence je svolení k činnosti, která je jinak zakázána. Licenční smlouva je samostatný smluvní typ, kterým poskytovatel uděluje nabyvateli svolení (licenci) k výkonu svých absolutních práv k nemotnému statku. (Boháček, 2008).

4.3.2.1 Creative Commons licence

Creative Commons (CC) je soubor licencí vytvořený odborníky pod hlavičkou Americké neziskové organizace Creative Common. První koncept byl představen v roce 2011. Iniciativa reaguje na současnou situaci v autorském právu a umožňuje autorům rychlé a jednoduché sdílení svých děl a to bez ztráty autorství. Vznik licence byl inspirován úspěchem softwarové licence GNU vyvinuté organizací Free Software Foundation. „Pod CC licencí lze zpřístupnit a poskytovat jakákoliv autorská díla nebo jiné nemotné statky, které jsou chráněny autorským zákonem. Přestože se CC

licence

v převážné míře používají na díla v elektronické podobě (především v síti Internet), je možné je aplikovat i na díla vyjádřená v hmotné podobě (např. tištěné dokumenty).“ (Gruber, 2011) (Creative Commons, 2017) Organizace CC je iniciátorem dalších projektů, zaměřených především na otevřené licencování v oblasti vědy a vzdělávání.

4.3.2.2 Licence Creative Archive

Licence Creative Archive je autorskou licencí pro publikaci archivních materiálů. Licence byla vytvořena skupinou Creative Archive Licence Group (CALG) v roce 2005. Mezi základní principy licence patří:

- Díla lze využít pouze pro nekomerční účely.
- Odvozená díla lze šířit pouze pod stejnou licencí.
- Je nutné vždy uvést autora původního díla.
- Dílo smí být využito jen ke kreativním účelům. Je zakázáno využívat dílo pro politické či charitativní kampaně. (BBC, 2016)

4.3.2.3 GNU Free Documentation License

GNU Free Documentation License (GNUFDL) je licence z roku 2000, vytvořená k licencování dokumentace softwaru. Licence byla vyvinuta organizací Free Software Foundation (FSF) jako doplněk k open source softwarové licenci GNU (GNU's Not Unix!). Princip licence vychází z jejího předchůdce pro sdílení obsahu Open Content License (Open Definition, 2017).

4.3.3 Otevřený přístup

Otevřený přístup (Open Access) je modelem vědecké komunikace, který zajišťuje trvalý, okamžitý a bezplatný přístup k výsledkům vědy a výzkumu na Internetu. Otevřený přístup vytváří alternativu k tradičnímu šíření vědeckých poznatků prostřednictvím vědeckých časopisů. Principem je neomezený online přístup k vědeckým informacím, který zajišťuje auto-

archivace článků v otevřených repositářích nebo publikování v otevřených časopisech.

Přínosy otevřeného přístupu uváděný zdrojem (Open Access, 2016)

- Zrychlení výměny vědeckých informací.
- Rozšíření dostupnosti vědeckých informací.
- Zvýšení viditelnosti vědeckých informací.
- Zvýšení čtenářské základny.
- Zvýšení informačního dopadu.

Jsou definovány dvě cesty otevřeného přístupu:

- Zlatá (Gold road) – publikování v otevřených časopisech, kde přístup poskytují vydavatelé.
- Zelená (Green road) – auto-archivace autorem v otevřených repositářích, kde přístup poskytují autoři. (Budapest Open Access Initiative, 2016)

4.3.4 Shrnutí

Významným trendem v současném světě je publikování informací pod licencemi, umožňující jejich rychlé a snadné šíření. V současné době patří mezi nejvhodnější licence pro šíření digitálních dokumentů soubor licencí, vyvinutý organizací *Creative Common*.

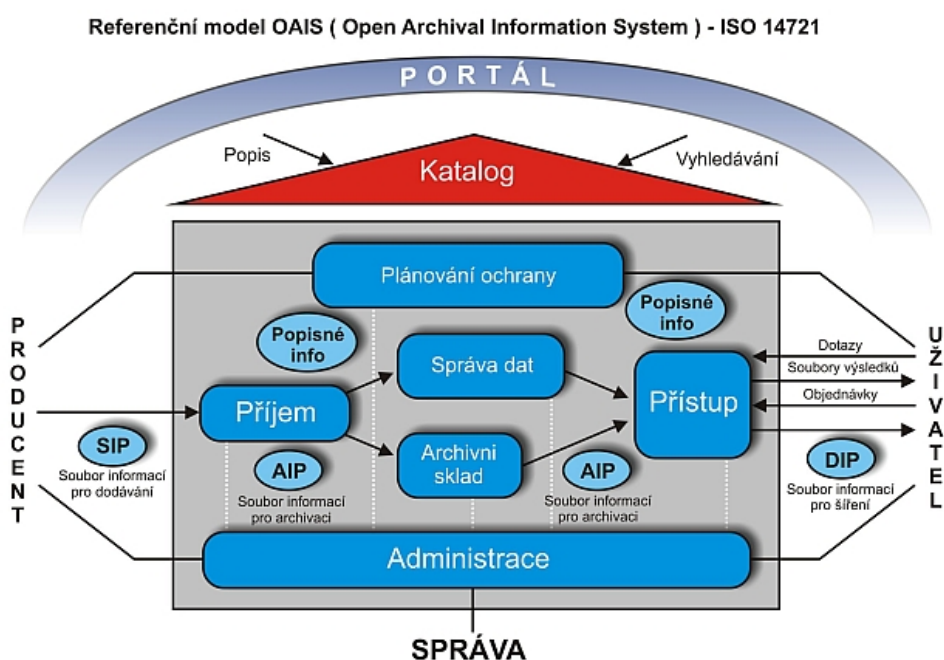
4.4 Aplikace pro podporu vědy

4.4.1 Digitální knihovny

Pojem digitální knihovna je velmi široký, definice je nejednotná. V odborné literatuře se vyskytují následující definice: „Spravovaná sbírka informací spolu s odpovídajícími službami, přičemž informace jsou uloženy v digitální podobě a jsou dostupné prostřednictvím sítě. (ARMS, 2000) „Integrovaný systém, zahrnující soubor elektronických informačních zdrojů a služeb umožňující získávání, zpracovávání, vyhledávání a využívání informací v tomto systému uložených. Digitální knihovny jsou zpřístupňovány prostřednictvím počítačových sítí. Účelem budování digitální knihovny je poskytnout uživatelům možnost jednotného přístupu k digitálním anebo digitalizovaným dokumentům, případně i k sekundárním informacím o tištěných primárních zdrojích, uložených ve fondu knihovny.“ (KTD, 2003) „Organizace, které poskytují zdroje (včetně specializovaného personálu) umožňující provádět výběr, strukturování a zpřístupnění sbírek digitálních prací, tyto práce dále distribuovat, udržovat jejich integritu a dlouhodobě uchovávat – a to vše s ohledem na snadné a ekonomické využití určitou komunitou nebo množinou komunit uživatelů” (Van de Sompel, 1999)

Mezi nejpropracovanější obecnou architekturu digitálních knihoven patří architektura dle Kahna a Wilenskeho (Kahn, 1995).

S pojmem digitální knihovna úzce souvisí pojem digitální repositář. Dle některých autorů není rozdíl mezi těmito pojmy, někteří autoři spojují pojem digitální repositář s konkrétními institucemi a principem otevřeného sdílení těchto dat. Problematikou digitálních repositářů se zabývá model OAIS (Open Archival Information System), který byl přijat jako norma (ISO 14721:2003). Princip modelu OAIP je znázorněn na obrázku 6.



Obrázek 12 – Referenční model OAIS. Zdroj: (Stoklasová, 2006).

Repositáře, které usilují o dlouhodobé uložení a zpřístupnění digitálních informací, usilují o statut důvěryhodného dlouhodobého repositáře. ISO 114721:2003, resp. 2012 je referenční model OAIS, norma, která definuje činnosti dlouhodobého repositáře, jeho cíle, a zavádí základní terminologii a informační model. ISO 14721:2003 definuje, jaká metadata má dlouhodobý repositář uchovávat. ISO norma 16363:2012 (navazující na Trusted repository audit checklist) je prostředkem pro certifikaci důvěryhodného dlouhodobého repositáře. Repositáře, které neukládají metadata vyžadovaná OAIS, a které nezveřejňují dokumentaci vyžadovanou ISO 16363 nelze z dlouhodobého hlediska považovat za důvěryhodné repositáře. (Cybr, 2003) (Planková, 2008)

Institucionální repositář

Institucionální repositáře, též lokální repositáře, shromažďují digitální objekty, které byly vytvořeny v rámci instituce, která repositář zřizuje. Vznikají hlavně na univerzitách a ve výzkumných institucích. Tematicky je obsah omezen pouze zaměřením zřizující instituce. (ÚK ČVUT, 2016)

COAR

The Confederation of Open Access Repositories (COAR) uvádí ve svém manifestu, že "Každý jednotlivý pramen má omezenou hodnotu pro výzkum: skutečná síla otevřeného přístupu spočívá v možnosti společného svázání všech úložišť". (Otevřený přístup, 2016)

Centrální repositář

Centrální repositáře nebo též předmětové repositáře jsou zaměřeny na určitou vědní oblast nebo obor. V rámci tohoto zaměření pak shromažďují (agregují) dokumenty z více institucí nebo i od samostatných vědců. Centrální repositáře poskytují vyhledávací služby nad metadaty, získanými z různých lokálních repositářů. (Müller, 2002)

Mezi centrální repositáře jsou zahrnuty repositáře, které agregují data dle určitého typu. Mezi tyto repositáře patří například:

- Repositář „šedé literatury“ Národního úložiště šedé literatury (NUŠL).
- Repositář COncecting REpositories (CORE) agregující stovky repositářů s otevřeným přístupem z různých zemí.
- Archív Europeana obsahující naskenovaná umělecká díla, filmy a knihy.
- Repositář dblp agregující metadata článků a příspěvků z oblasti vědy o počítačích (computer science) <http://dblp.uni-trier.de/>.

4.4.2 Aplikace využívající metadata

Metadata digitálních artefaktů (objektů) z digitálních knihoven jsou využívána aplikacemi pro podporu vědy. Jedná se především o webové aplikace, které se dělí do následujících skupin:

- sociální sítě pro vědecké pracovníky (Social network service for scientists),
- aplikace pro zpravu referencí (reference management software),
- vyhledávače odborných vědeckých prací (web search engine of scientific work).

4.4.2.1 Sociální sítě pro vědecké pracovníky

Sociální sítě pro vědecké pracovníky (Social network service for scientists) jsou sociální sítě vyvinuté pro vědecké pracovníky a slouží k podpoře jejich činnosti – především podpoře vzájemné komunikace a sdílení znalostí. Zejména mladší uživatelé využívají ke komunikaci a sdílení svých znalostí sociální média. Dle Stočese (2015) je doporučeno některé funkce sociálních sítí integrovat přímo do systémů na řízení výuky (LMS) (learning management systems), které zde pak slouží jako komunikační nástroj mezi přednášejícími a studenty či dokonce studenty mezi sebou. A umožní získat studentů nejnovější poznatky z dané oblasti.

ResearchGate

ResearchGate je sociální síť, speciálně vytvořená pro vědeckou komunitu. Platforma má uživatelům usnadnit vzájemnou komunikaci a spolupráci a umožnit bezplatný přístup k vědecké literatuře. ResearchGate nabízí vytvoření osobního profilu, členství v zájmových skupinách, využití burzy práce se specializací na pozice pro vědce a akademiky. Platformy také poskytují informace o aktuálních vědecky orientovaných událostech a umožňují přístup k vědeckým databankám. (Kintisch, 2014)

academia.edu

Academia.edu je sociální síť pro akademiky. Plošně může být použita ke sdílení dokumentů, sledovat jejich dopad a sledovat výzkum v daném oboru. Její provoz byl zahájen v září 2008 [6], s 31 miliony registrovaných uživatelů v lednu 2016 a více než 8 miliony nahranými příspěvky. Academia.edu založil Richard Price.⁴

⁴ <https://www.academia.edu/about>

Velkým trendem současnosti je vznik oborově zaměřených sociálních sítí. Oborově zaměřené sociální sítě lze definovat jako platformu pro online sdílení informací osob se stejným profesním zaměřením.

VOA3R

VOA3R, neboli Virtual Open Acces Agriculture & Aquaculture Repository, je evropský projekt, jehož se účastní 14 subjektů z 10 evropských zemí. Hlavním cílem projektu je realizace otevřeného přístupu do zemědělského a vodohospodářského archivu, který obsahuje vědecké a odborně-výzkumné dokumenty, vztahující se k zemědělství, potravinářství a životnímu prostředí. Cílovými koncovými uživateli realizovaného systému jsou vědečtí pracovníci a akademici, odborníci z oblasti zemědělství, potravinářství, vodohospodářství, životního prostředí a venkovských regionů, a studenti (zejména studenti postgraduálního studia). Cílová skupina, tedy vědecko-výzkumní pracovníci, mají pomocí tohoto archivu získat jednotné místo informační podpory pro jejich obor. (Šimek, 2012) (Sgouropoulou, 2011)

Dalšími platformami, zaměřenými na vědeckou komunitu jsou například *MyScienceWork* nebo *ScienceStage*.

4.4.2.2 Aplikace pro zprávu referencí

Aplikace pro zprávu referencí (reference management software) jsou aplikace sloužící ke správě referencí nebo osobnímu bibliografickému managementu.

Tyto softwarové balíky se obvykle skládají z databáze, ve které lze zadat úplné bibliografické odkazy a dále také systém pro generování selektivních seznamů článků v různých formátech, které vyžadují vydavatelé a vědecké časopisy. Moderní balíčky správy odkazů mohou být obvykle integrovány s textovými procesory, takže seznam článků v příslušném formátu se automaticky vytvoří podle článku, čímž se sníží riziko, že citovaný

zdroj nebude zařazen do seznamu odkazů. V rámci těchto systémů lze propojovat metadata s konkrétními profily autorů. Jedná se především o komerční aplikace. Mezi nejvýznamnější aplikace tohoto typu patří:

- **Mendeley**⁵ provozován společností Elsevier - provozovatelem citační databáze SCOPUS
- **EndNote** ⁶ provozován společností Thomson Reuters - provozovatelem WoS
- **REFWORKS**⁷

ORCID

Do této skupiny lze zařadit i aplikace identifikátoru ORCID (Open Research and Contributor ID), pomocí které lze zpravovat bibliografické zápisy osobních publikací.

4.4.2.3 Vyhledavače odborných vědeckých prací

Existuje celá řada vyhledavačů odborné literatury, většina z nich je obrově zaměřená.

scholar.google.cz⁸

Google Scholar je volně přístupný webový multidisciplinární vyhledavač, který indexuje plný text nebo metadata odborné literatury v celé řadě publikačních formátů. Index Google Scholar, vydaný v beta verzi v listopadu 2004, obsahuje nejvíce recenzovaných on-line akademických časopisů a knih, konferenčních článků, tezí a disertací, předtisků, abstraktů, technických zpráv a jiné odborné literatury, včetně soudních posudků a patentů. Google scholar je nejrozšířenější a nejobsáhlejší aplikací v dané kategorii.

Google Scholar umožňuje propojení vlastního google profilu s indexovanými články.

⁵ <https://www.mendeley.com/>

⁶ <https://projectne.thomsonreuters.com/#/login?app=endnote>

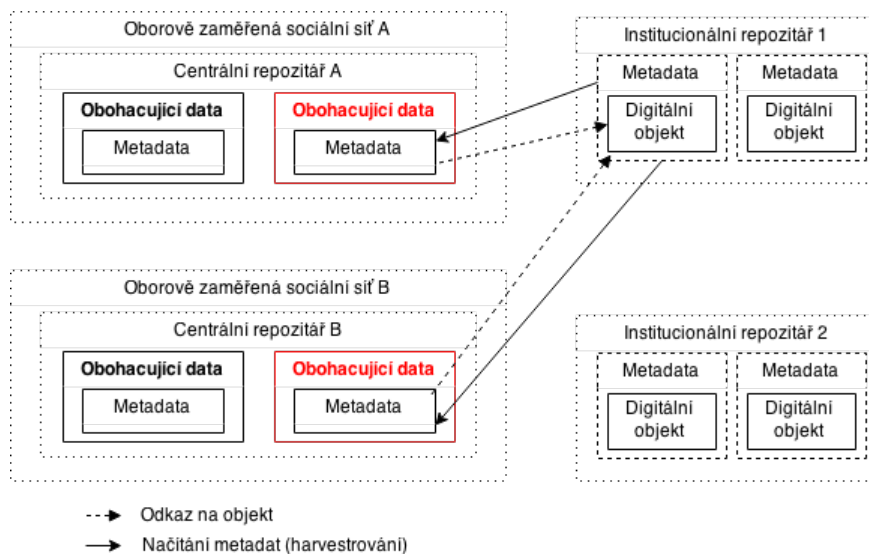
⁷ <https://www.refworks.com/>

⁸ <https://scholar.google.cz/>

4.4.3 Shrnutí

Aplikace typu sociální síť pro vědecké pracovníky umožňují uživatelům vyhledávat objekty z repositářů a přidávat k nim další data. Tato data lze nazývat daty „obohacujícími“. Obohacující data jsou uložena v rámci sociální sítě a může být s nimi pracováno pouze prostřednictvím té konkrétní sítě. Princip je znázorněn na obrázku Obrázek 13. Obohacující data lze klasifikovat do dvou skupin, a to *vazebních* a *dalších* metadat. Strukturu a funkci *dalších metadat* si vytváří každá sociální síť zvlášť. Příkladem těchto dat mohou být například komentáře, hodnocení (rating) apod. Mezi vazební metadata patří vazby:

- Digitální artefakt <-> Osoba (autor, spoluvůrci)
- Digitální artefakt <-> Digitální artefakt (citace, reference)



Obrázek 13 – vztah institucionálních repositářů a oborově zaměřených soc. sítí. Zdroj: Autor.

4.5 Poskytování metadat pro podporu vědy

4.5.1 Analýza a zhodnocení vybraných institucionálních repositářů

Následující kapitola se zaměřuje na analýzu stavu vybraných institucionálních repositářů. Pro výzkum byly vybrány české institucionální repositáře indexované v databázi Web of science za rok 2016. Byly analyzovány aspekty otevřenosti repositáře a dále problematika poskytování metadat o digitálních artefaktech uložených v lokálních repositářích pomocí veřejného rozhraní. Podklady pro analýzu byly získávány z veřejně dostupných informací prezentovaných na oficiálních stránkách daného vědeckého časopisu. Podrobné výsledky se nacházejí v tabulce 6 v příloze F.

Celkem bylo analyzováno 49 vědeckých časopisů. Bylo zjištěno, že žádný z vybraných zkoumaných repositářů neposkytuje svá data pomocí veřejného API. Některé vědecké časopisy k publikaci svých článků využívají služeb vydavatele Springer Publishing ⁹, který poskytuje metadata publikovaných článků jako placenou službu. Oproti tomu časopisy poskytují svá metadata pomocí protokolu OAI-PMH knihovnám např. Knihovna AV ČR, v. v. i. ¹⁰. Všechny repositáře poskytují vyhledávací funkce, z čehož lze usuzovat, že publikované články popisují metadata a disponují nějakým aplikačním protokolem.

Třicet šest z hodnocených repositářů poskytují obsah bez poplatku a jsou licencované některou z otevřených licencí, lze je tedy zařadit do kategorie otevřené časopisy (Open access journal). Dva medicínské časopisy sice poskytují bez poplatku celé obsahy (full paper) článků, ale přístup jen po odsouhlasení podmínky, že čtenář je lékařem nebo odborníkem v oblasti zaměření časopisu.

⁹ <http://www.springer.com/>

¹⁰ <https://www.lib.cas.cz/>

4.5.2 Identifikace digitálního artefaktu

4.5.2.1 Formáty bibliografické citace

Existuje celá řada formátů určených pro zápis bibliografických citací. Tyto formáty jsou spjaty s konkrétními SW řešeními. Mezi nejrozšířenější patří BibTeX a EndNote. Dále existuje například pracovní skupina při Dublin Core zabývající se problematikou zápisu bibliografických citací Dublin Corem.

BibTeX

BibTeX je referenční software pro formátování seznamů odkazů. Nástroj BibTeX se obvykle používá společně s přístrojem pro přípravu dokumentů LaTeX. Cílem BibTeXu je snadné uvést zdroje konzistentně, oddělením bibliografických informací od prezentace těchto informací, podobně jako oddělení obsahu a prezentace / stylu podporovaných samotným LaTeXem. Ukázka zápisu:

```
@misc{ Nemcova01,  
      author = " Němcová Božena ",  
      title = "Babička",  
      year = "1855" }
```

EndNote

EndNote je komerční softwarové řešení pro správu referencí, který slouží ke správě bibliografií a odkazů při psaní dokumentů. EndNode má svůj formát zápisu citací. Níže je uvedena ukázka zápisu:

```
%0 Book  
%A Němcová Božena  
%D 1855  
%T Babička
```

Identifikátory

DOI

DOI (Digital Object Identifier) (identifikátor digitálního objektu) je centralizovaný komerční systém identifikátorů děl přístupných v digitální podobě. DOI je popsán standardem ISO 26324. DOI zabezpečuje jednoznačnou identifikaci digitálního dokumentu na internetu, je neměnný a poskytuje trvalý odkaz na dokument. V současné době je identifikátor DOI standardem pro identifikaci vědeckých publikací.

ISBN

ISBN (International Standard Book Number) (mezinárodní standardní číslo knihy) je jednonačný identifikátor knihy. Pro upřesnění identifikace konkrétního článku je nutné udávat i číslo stránky.

ISSN

ISSN (International Standard Serial Number) (mezinárodní standardní číslo seriálové publikace) je číslo, určeno pro periodika. Pro upřesnění konkrétního článku je vhodné jej doplnit o pořadové číslo, ročník, stránku.

Další identifikátory

- EID₂ – identifikátor v citační databázi SCOPUS
- WOS Accession Number – identifikátor v citační databázi WoS
- ASIN – Amazon Standard Identification Number
- ISAN – International Standard Audiovisual Number
- ISMN – International Standard Music Number for Printed Music
- ISRC – The International Standard Recording Code
- ISRN – International Standard Technical Report Number
- ISWC – International Standard Musical Work Code

4.5.3 Identifikace osob

Příspěvky konkrétního autora vědecké publikace mohou být těžko rozpoznatelné, protože většina osobních jmen není jedinečná. Další nedostatky zápisu jmen jsou následující:

- mohou se měnit (například s manželstvím),
- existují kulturní rozdíly v pořadí jmen,
- není sjednoceno použití zkratk pro jména,
- existuje celá řada systémů zápisu.

K přiřazení konkrétního autora (jeho profilu) k článku nedochází při publikaci, ale až v systémech na ně navázaných. Existuje řada systémů identifikace autorů, které se snaží nedostatky odstranit.

ResearchID

Research ID je systém identifikace vědeckých autorů. Systém byl zaveden v lednu 2008 společností Thomson Reuters. Je využíván ve stejnojmenném systému a v citační databázi WoS.

Scopus Author ID

Scopus Autor ID pomáhá se správou bibliografických citací v aplikaci Mendeley a citační databázi SCOPUS.

ORCID

ORCID (*Open Research and Contributor ID*) je neproprietární alfanumerický kód, který jednoznačně identifikuje vědecké a jiné akademické autory a přispěvatele. Poskytuje lidem trvalou identitu podobnou té, která byla vytvořena pro subjekty související s obsahem v digitálních sítích pomocí identifikátorů digitálních objektů (DOI). ORCID se snaží o spojení identifikátoru ResearchID a Scopus Autor ID.(viz Obrázek 14) ORCID je spravován neziskovou organizací. (Nature, 2009)



Obrázek 14 – ORCID - výpis ID různých systémů. Zdroj: Autor.

Dalším přístupem pro identifikaci autora je jeho email. Každý email patří k jedné osobě a dá se z něj většinou vyčíst jméno a organizaci daného uživatele. Vhodné je doplnit email o datum, ke kterému je platný.

Mezi další systém pro identifikaci autora patří například: The Digital Author Identifier (DAI).

4.5.4 Analýza sociálních sítí pro vědecké pracovníky

V následujícím oddíle jsou analyzovány vybrané aplikace typu sociální síť pro vědecké pracovníky. Systémy jsou analyzovány (viz Tabulka 6) na základě následujících aspektů:

- pomocí jakého identifikátoru je ztotožněn autor,
- pomocí jakého identifikátoru je určen digitální artefakt,
- možnost importovat metadata o digitálním artefaktu
- možnost exportovat metadata o digitálním artefaktu

Název	ID autor	ID objekt	Import	Export
Mendeley	Scopus autor ID, ORCID	DOI a další	ano různé	ano různé
ResearchID/EndNote	ORCID, researchID	DOI a další	ano různé	ano různé
Google Scholar	vlastní	vlastní	ne	ano různé
ResearchGate	vlastní	DOI, vlastní	ano různé	ne
Academia edu	vlastní	vlastní	ne	ne
ORCID	ORCID, Scopus autor ID, researchID	Dle databáze importu.	ano různé	ano BibTex

Tabulka 6 – Analýza aplikací pro podporu vědy. Zdroj: Autor.

Sociální sítě pro vědecké pracovníky neumožňují uživatelům si data v nich vytvořená exportovat k identifikaci autora nepoužívají své vlastní proprietární identifikátory.

Většina analyzovaných systémů umožňuje export záznamů, ale neobohacuje je o žádná dodatečná data např. identifikaci autora pomocí ORCID.

ORCID DOI jsou používané identifikátory většiny analyzovaných aplikací.

4.5.5 Shrnutí

Nejednotnost identifikace autora je stále problémem převážně v systémech, které publikují digitální artefakty. Identifikátory, které jsou čím dál více rozšiřovány a využívány jsou ORCID – identifikace autorů, DOI –

identifikace digitálních objektů. ORCID se snaží spojit dva proprietární identifikátory researchID a Scopus Autor ID.

Pro navrhovanou metodiku se doporučuje uvést všechny dostupné identifikátory, zkvalitní to výsledný záznam.

4.6 Metadatové modely v praxi

Jako vhodné metadatové modely pro popis metadat vzniklých činností sociálních sítí pro vědecké pracovníky dle předchozích oddílů patří LOM, DC, MODS. V následujícím oddíle bude rozebrána problematika popisu vazeb na autora a jiné digitální artefakty vybraných modelů.

K zápisu metadat v následujících pododdílech byl použit datový formát XML.

4.6.1 Vazba a identifikace autora

Schéma Dublin Core

Autoři jsou identifikováni elementem *dc:creator*, který dle standardu neobsahuje další rozšíření.

```
<dc:creator>Michal Stočes</dc:creator>
```

Datový model LOM

Problematika vazby na osobu je v modelu LOM řešena velmi komplexně a pomocí elementu *lom:lifeCycle* a jeho potomků. Ve standardu chybí možnost detailnější identifikace autora.

```
<lom:lifeCycle>
  <lom:contribute >
    <lom:role>
      <lom:source>LOMv1.0</lom:source>
      <lom:value>Author</lom:value>
    </lom:role>
    <lom:entite>
      Michal Stočes
    </lom:entite>
  </lom:contribute >
</lom:lifeCycle>
```


Schéma MODS

```
<mods:name type="personal">
  <mods:namePart type="family">Michal</mods:namePart>
  <mods:namePart type="given">Stořes</mods:namePart>
  <mods:role>
    <mods:roleTerm type="code" authority="marcrelator">
      aut
    </mods:roleTerm>
    <mods:roleTerm type="text" authority="marcrelator">
      Author
    </mods:roleTerm>
  </mods:role>
</mods:name>
```

4.6.2 Vazba a identifikace reference, citace

Ve všech třech modelech je rozdíl mezi referencí a citací tvořen kvantifikátorem `references/isReferedBy`,

Schéma Dublin Core

Jako zápis metadat identifikující referenci je zde použit kvalifikovaný zápis Dublin Core. Ve standardu DC neexistuje jednoznačný postup, jak zapsat odkaz na konkrétní místo v dokumentu (stránky).

```
<dcterms:references xsi:type="dcterms:URI">
  http://doi:10.17221/313/2015
</dcterms:references >
```

Datový model LOM

Standard IEEE LOM má stejné nedostatky jako DC - neexistuje jednoznačný postup, jak zapsat odkaz na konkrétní místo v dokumentu (stránky).

```
<lom:relation>
  <lom:kind>
    <lom:source>LOMv1.0</lom:source>
    <lom:value>references</lom:value>
  </lom:kind>
<lom:resource>
```

```

    <lom:identifier>
      <lom:catalog>URI</lom:catalog>
      <lom:entry>
        http://dx.doi.org/10.7160/aol.2016.080108
      </lom:entry>
    </lom:identifier>
  </lom:resource>
</lom:relation>

```

Schéma MODS

Schéma MODS obsahuje v definici popis umožňující odkaz na konkrétní místo v dokumentu. Dále element `mods:identifier` má vlastnost `type`, u které není určeno, jakých hodnot může nabývat, což dělá MODS velmi univerzálním z pohledu jednoznačné identifikace díla.

```

<mods:relatedItem type="references">
  <mods:identifier type="doi">
    10.7160/aol.2016.080108
  </mods:identifier>
  <mods:part>
    < mods:extent unit="pages">
      <mods:start>85</mods:start>
      <mods:end>86</mods:end>
    </mods:extent>
  </mods:part>
</mods:relatedItem>

```

4.6.3 Mapování metadatových modelů

Následující tabulky (7, 8) demonstrují mapování jednotlivých metadatových modelů.

Dublin Core (DC) elementy	Learning object metadata (LOM) elementy
dc:identifier	/lom/general/identifier/entry
dc:title	/lom/general/title
dc:language	/lom/general/language
dc:description	/lom/general/description

dc:subject	/lom/general/keyword <i>nebo</i> /lom/classification s /lom/classification/purpose <i>je rovno</i> "discipline" <i>nebo</i> "idea".
dc:coverage	/lom/general/coverage
dc:type	/lom/educational/learningResourceType
dc:date	/lom/lifeCycle/contribute/date <i>když</i> /lom/lifeCycle/contribute/role <i>nabývá hodnoty</i> "publisher".
dc:creator	/lom/lifeCycle/contribute/entity <i>když</i> /lom/lifeCycle/contribute/role <i>nabývá hodnoty</i> "author".
dc:otherContributor	/lom/lifeCycle/contribute/entity <i>s typem příspěvku uvedeným v</i> /lom/lifeCycle/contribute/role
dc:publisher	/lom/lifeCycle/contribute/entity <i>když</i> /lom/lifeCycle/contribute/role <i>nabývá hodnoty</i> "publisher".
dc:format	/lom/technical/format
dc:rights	/lom/rights/description
dc:relation	/lom/relation/resource/description
dc:source	/lom/relation/resource <i>když</i> <i>nabývá hodnoty</i> /lom/relation/kind is "isBasedOn".

Tabulka 7 - mapování mezi DC a LOM

MODS elementy	DC elementy
<titleInfo><title>	Title
<name><namePart>	Creator, Contributor
<subject> <topic> <classification> <name> <occupation>	Subject
<abstract> <note> <tableOfContents>	Description
<originInfo><publisher>	Publisher
<originInfo><dateIssued> <originInfo><dateCreated> <originInfo><dateCaptured> <originInfo><dateOther>	Date
<typeOfResource><genre>	Type

<physicalDescription><internetMediaType> <extent><form>	Format
<identifier><location> <url>	Identifier
<language><languageTerm>	Language
<relatedItem>	Relation
<subject> <geographic> <temporal> <hierarchicalGeographic> <cartographics>	Coverage
<accessCondition>	Rights

Tabulka 8 - mapování mezi DC a MODS

(Ehlers, 2006)

4.6.4 Shrnutí

Vybrané metadatové modely DC, LOM a MODS byly analyzovány a byly zjištěny následující závěry a doporučení pro navrhovaný metodický postup:

- DC je vhodný formát pro zápis základních metadat.
- Pro popis vazeb na osoby je vhodný standard LOM doplněný elementem identifier z modelu MODS.
- Pro popis vazeb na digitální artefakt je vhodný model MODS.
- Popis dalších obohacujících dat je nutno vytvořit vzhledem k jejich povaze.
- Schémata lze mezi sebou mapovat / transformovat.

4.7 Hypotéza disertační práce

Na základě syntézy získaných poznatků bylo zjištěno, že dosud nebyl publikován metodický postup pro sdílení metadat aplikací typu sociální síť pro vědecké pracovníky. Proto je dále navržena metodika, která má za cíl zlepšit podmínky pro práci v komunitě vědeckých pracovníků.

5 Návrh nové metodiky

Na základě poznatků uvedených v předchozím oddíle je navrhován nový metodický postup pro zkvalitnění přenosu metadat, která jsou obohacena (data enrichment) činnostmi aplikací typu sociální síť pro vědecké pracovníky (Social network service for scientists).

V následujících oddílech disertační práce je navrhován metodický postup označovaný jako MSOD (metodika sdílení obohacených dat). K prezentaci návrhu je využíván diagram aktivit dle specifikace UML (Unified Modeling Language).

5.1 Předpoklady pro použití metodiky

Navržený metodický postup je určen pro provozovatele aplikací typu sociální síť pro vědecké pracovníky (Social network service for scientists). Jeho použitím je vytvořen aplikační profil zápisu metadat digitálního artefaktu (objektu), vyskytujícího se v dané aplikaci. Aplikační profil umožní uživatelům aplikací typu sociální síť pro vědecké pracovníky importovat nebo exportovat do/z aplikace metadata popisující digitální artefakt (objekt).

5.2 Formulace metodiky MSOD

Navrhovaná metodika pro přenos obohacujících dat se skládá ze dvou etap.:

1. Identifikace metadat popisujících digitální artefakty (objekty).
2. Tvorba aplikačního profilu.

5.2.1 Etapa 1 Identifikace metadat

Přípravná etapa je zaměřena na identifikaci a kategorizaci struktury datového modelu (metadatové struktury) aplikace, ke kterému budou v následující etapě přiřazeny konkrétní elementy. Etapa 1 je tvořena dvěma fázemi. Identifikace datového modelu (metadatové struktury) vychází z poznatků získaných v oddíle 4.4.

Fáze 1 Identifikace primárních metadat

V první fázi jsou identifikována metadata, která nevznikla činností aplikací typu sociální síť pro vědecké pracovníky. Jedná se o metadatové popisy vzniklé při publikaci digitálních artefaktů. Např: *název, autoři, licence* apod. Tyto elementy jsou dále nazývány termínem primární elementy. Každý lokální repositář (vydavatel) má jiný metadatový model, proto je nutné primárních elementy identifikovat jako průnik všech metadatových modelů jehož digitální artefakty se vyskytují v aplikaci.

Fáze 2 Identifikace obohacujících dat

Ve druhé fázi jsou identifikována metadata, která vznikla činností aplikací typu sociální síť pro vědecké pracovníky. Nejprve jsou identifikována metadata, sloužící k identifikaci vazeb na profily autorů nebo jiné digitální artefakty (citace, reference, projekty). Dále následuje identifikace dalších obohacujících dat, specifických ke konkrétní aplikaci, příkladem těchto dat mohou být různé komentáře, rating (hodnocení), apod.

5.2.2 Etapa 2 Tvorba aplikačního profilu

Etapa 2 se skládá ze sedmi fází, vycházejících z poznatků a postupů CEN/ISSS a singapurského frameworku pro tvorbu aplikačních profilů (viz oddíl 4.2.6). K popisu jsou využívány elementy jmenného prostoru metadatových modelů DC, MODS a LOM (viz 4.6 a 4.2).

Závislost jednotlivých fází je znázorněna v diagramu aktivit dle specifikace UML (viz Obrázek 16), pro aktivitu je v metodice použit termín fáze.

Výsledný aplikační profil může být prezentován ve formátu tabulky, popisující jednotlivé elementy a jejich vlastnosti, nebo jako zápis XML šablony ve formátu XSD nebo RDF.

Fáze 1. Výběr datových elementů pro zápis primárních metadat

Nejprve budou vybrány elementy pro popis primárních (původních) metadat, ze kterých bude tvořen nový aplikační profil. K popisu budou vybrány elementy ze jmenného prostoru Dublin Core.

Fáze 2. Přidání nových lokálních elementů obohacujících dat

Budou přidány nové lokální elementy (elementy vytvořené činností sociální sítě) obohacujících dat. Tuto fázi lze rozdělit do čtyř kroků.

Výběr elementů identifikující digitální artefakt

Datová část identifikující samotný digitální artefakt je vytvářena pomocí standardu MODS s využitím elementu z *identifier*.

Výběr elementů identifikující vazbu na osoby

Datová část *identifikující vazbu na osoby* (vazba: *digitální artefakt <-> osoba*) je vytvořena kombinací elementů jmenných prostorů standardů LOM a MODS. Jedná se o elementy: *lom:lifeCycle*, *lom:contribute*, *lom:role*, *lom:source*, *lom:value*, *mods:identifier*. Mezi vazby na osoby patří především identifikace autora digitálního artefaktu.

Výběr elementů identifikující vazbu na jiné digitální artefakty

Datová část *identifikující vazbu na jiné digitální artefakty* (vazba: *digitální artefakt <-> jiné digitální artefakty*) je vytvořena pomocí schématu MODS. Jsou použity následující elementy: *relatedItem identifier*. Mezi vazby na další digitální artefakty patří především identifikace citací a referencí.

Výběr elementů vhodných pro zápis dalších metadat

Struktura popisu dalších metadat je navržena s ohledem na jejich povahu, v závislosti na konkrétní sociální síti, a to použitím elementů z některých ze standardů LOM, DC, MODS nebo jejich kombinací, nebo definováním nových elementů. Příkladem může být rozšíření klíčových slov

o položky z řízeného slovníku. Nebo obahuje-li sociální síť některé prvky vzdělávání, lze využít příslušné specifické elementy standardu LOM.

Fáze 3. Určení míry povinnosti elementů

V tomto kroku je určena míra povinnosti výskytu jednotlivých elementů.

- povinné (mandatory)
- doporučené (recommend)
- podmíněné (conditional)
- volitelné, nepovinné (optional)

Mezi povinné elementy musí patřit minimálně název digitálního artefaktu (<dc:title>) a typ (<dc:type>) – jedná li se o text, vizualizaci, zvuk apod.

Fáze 4. Určení min. a max. počtu výskytu elementů

Dále je zapotřebí určit minimální a maximální počet výskytu jednotlivých elementů. U elementů identifikující vazbu na další artefakty nebo osoby je doporučeno neomezovat horní hranici výskytu – viz závěry z oddíly 4.5.2 .

Fáze 5. Definice datových a výčtových typů

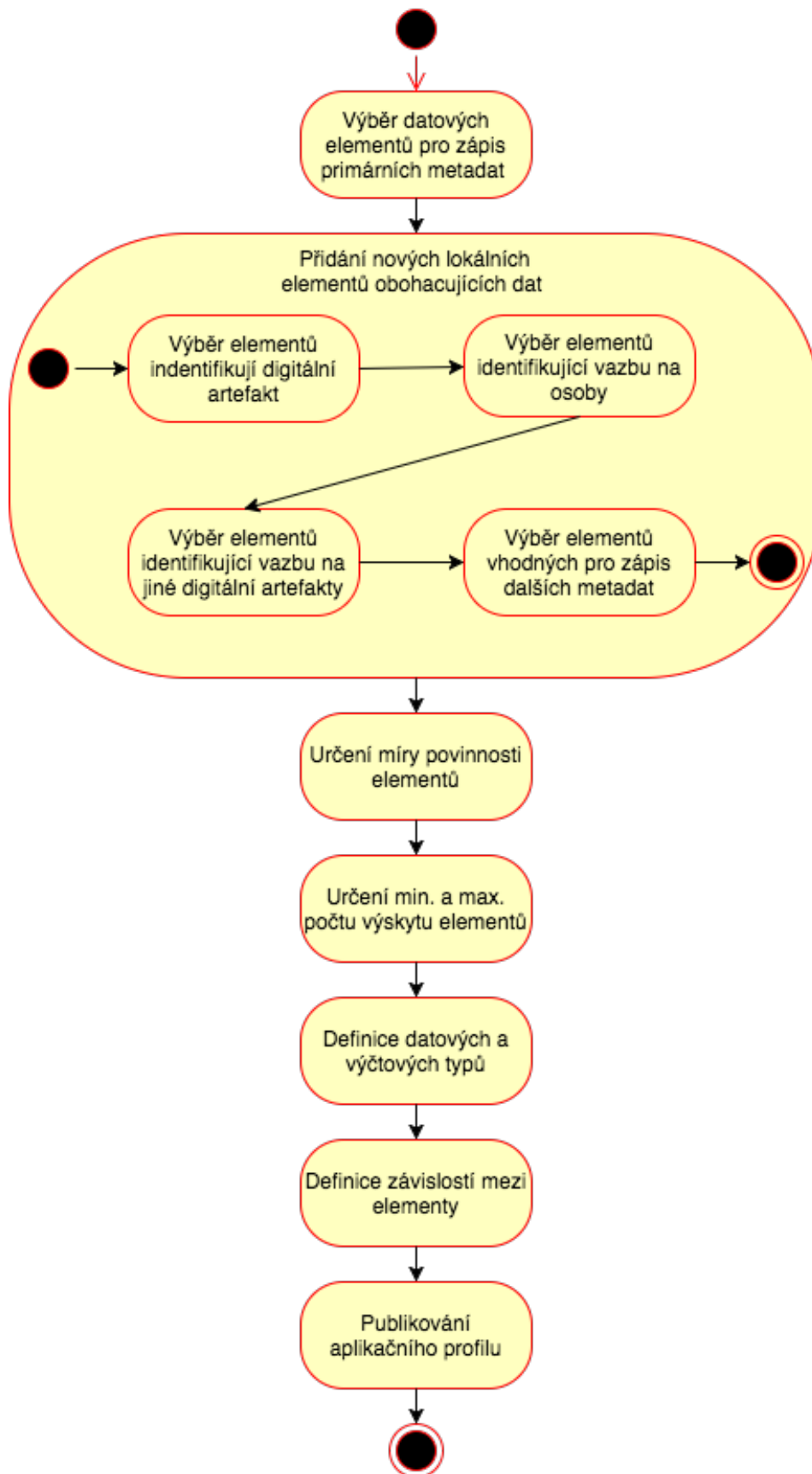
Definice datových typů a výčtů hodnot vychází z použitého standardu modelu metadat anebo lze vytvořit nový datový výčtový typ.

Fáze 6. Definice závislostí mezi elementy

Závěrečná fáze před publikací aplikačního profilu je definování závislostí mezi jednotlivými elementy. Definování závislostí slouží především k vyjádření logické závislosti některých elementů a také k redukci duplicitních zápisů.

Fáze 7. Publikování aplikačního profilu

Závěrečná fáze představuje vytvoření aplikačního profilu aplikace reprezentované jako XML schéma, RDF formát nebo textový seznam obsahující názvy elementů, jejich vlastnosti a omezení.

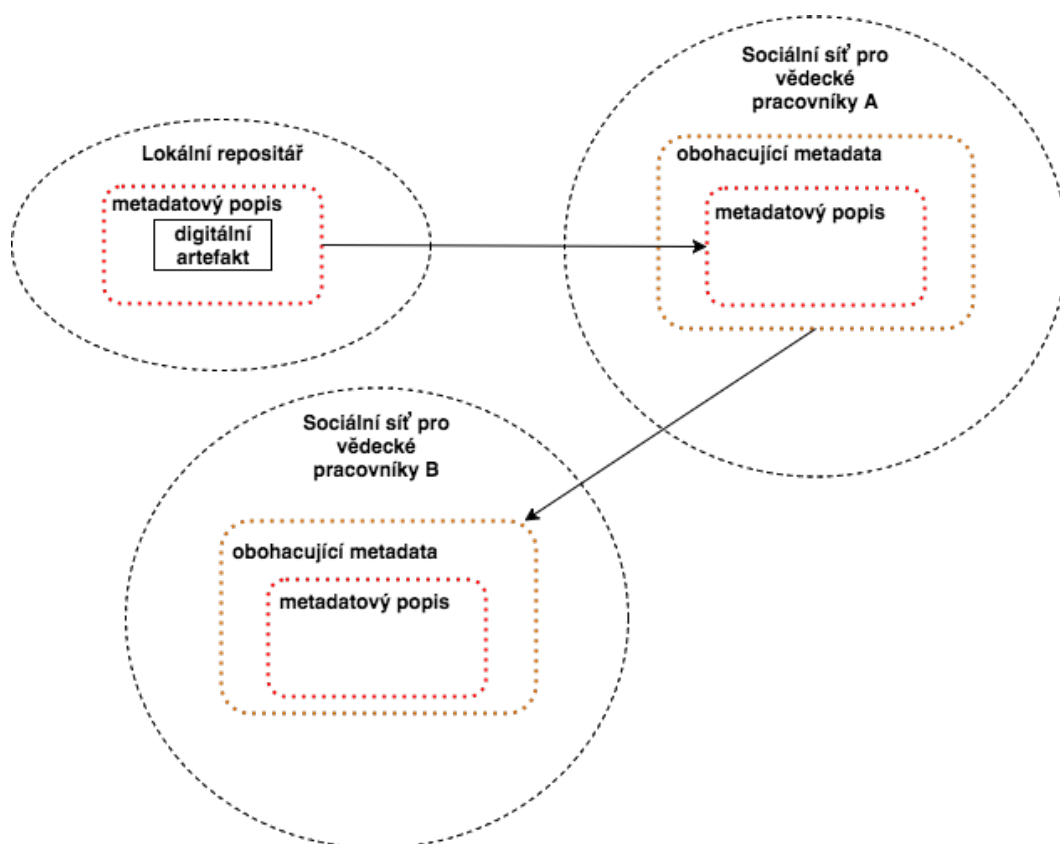


Obrázek 15 – MSOD - diagram aktivit tvorby aplikačního profilu.
Zdroj: Autor.

Navazující oddíl 6 této práce se zabývá ověřením metodiky, toto ověření slouží i k demonstraci praktického použití metodiky při tvorbě aplikačního profilu ve formátu XMD.

5.3 Očekávané přínosy navržené metodiky

Hlavním přínosem metodiky MSOD je zjednodušení přenosu metadatových popisů mezi aplikacemi typu sociální síť pro vědecké pracovníky. Obrázek 16 popisuje princip přenosu metadat z lokálního repozitáře do aplikace typu sociální síť pro vědecké pracovníky a následný přenos obohacujících metadat do jiné aplikace. K přenosu mezi lokálním repozitářem a aplikací může docházet automatickým načtením (např. pomocí protokolu OAI-PMH) nebo přepisem údajů uživatelem. Metadata z aplikace A jsou exportována ve formátu XML a následně jsou přenesena do aplikace B. Přenos může být realizován manuálně uživatelem anebo pomocí protokolu OAI-PMH.



Obrázek 16 – přenos metadat mezi aplikacemi. Zdroj: Autor

Použitím této metodiky vznikne aplikační profil, na základě kterého lze vytvořit konkrétní metadatové záznamy konkrétních digitálních artefaktů. Tyto artefakty lze dále snadno transformovat a využívat k přenosu dat mezi dalšími aplikacemi, sloužící pro podporu vědecké práce.

Očekává se, že výsledný záznam bude čitelný a vhodný k využívání lidmi i stroji (Human-readable, Machine-readable).

6 Ověření a zhodnocení metodiky MSOD

6.1 Sestavení aplikačního profilu

Následující oddíl demonstruje použití a ověření navržené metodiky MSOD. Bude vytvořen aplikační profil článku v sociální síti pro vědecké pracovníky. Aplikační profil bude vytvořen ve formě XML. Výsledné schéma se nachází v příloze G.

6.1.1 Etapa 1 Identifikace metadat

Fáze 1 identifikace primárních metadat

Pro ověření metodiky byly vybrány následující primární elementy:
název digitálního artefaktu, typ, datum publikace, souhrn.

Fáze 2 Identifikace obohacujících dat

Pro ověření metodiky byly vybrány následující obohacující datové popisy:
Identifikace, reference, citace, autoři, hodnocení, klíčová slova z tezauru AGROVOC.

6.1.2 Etapa 2 Tvorba aplikačního profilu

Fáze 1. Výběr datových elementů pro zápis primárních metadat

Pro primární elementy identifikované v první fázi první etapy byly ze standardu Dublin Core vybrány následující elementy:

- *název digitálního artefaktu - <dc:title>*,
- *typ - <dc: type>*,
- *datum publikace-<dc: date>* ,
- *souhrn -<dc:description>*.

Fáze 2. Přidání nových lokálních elementů obohacujících dat

Výběr elementů identifikující digitální artefakt

Dle metodického postupu je identifikace vytvářena pomocí standardu MODS s využitím elementu z *<mods:identifier>*.

Výběr elementů identifikující vazbu na osoby

Dle metodického postupu je autor identifikován kombinací elementů jmenných prostorů standardů LOM a MODS a to elementy: *<lom:lifeCycle>*, *<lom:contribute>*, *<lom:role>*, *<lom:source>*, *<lom:value>*, *<mods:identifier>*.

Výběr elementů identifikující vazbu na jiné digitální artefakty

Dle metodického postupu jsou vybrány elementy pro popis citací a referencí ze schématu MODS, a to konkrétně element *<mods:relatedItem>*, *<mods:identifier>*, *<mods:part>*, *<mods:extent>*, *<mods:start>* a *<mods:end>*.

Výběr elementů vhodných pro zápis dalších metadat

Dle metodiky MSOD je definován nový element *<rating>* sloužící k popisu hodnocení. Popis klíčovými slovy z tezauru AGROVOC je vybrán elementem *<dc:subject>* ze jmenného prostoru modelu Dublin Core.

Fáze 3. Určení míry povinnosti elementů

Je určeno, že atributy *<dc:title>*, *<dc:type>* budou povinné, ostatní ne.

Fáze 4. Určení min. a max. počtu výskytu elementů

Bylo rozhodnuto o výskytu elementů následovně:

- *<dc:title>* – výskyt právě jednou
- *<dc:type>* – výskyt právě jednou
- *<dc:date>* – výskyt právě jednou
- *<dc:description>* – neomezeně
- *<mods:identifier>* – neomezeně

- <lom:lifeCycle>- neomezeně
- <lom:contribute>- neomezeně
- <lom:role>- neomezeně
- <lom:source>- neomezeně
- <lom:value>- neomezeně
- <dc:subject>- neomezeně
- <mods:part>- neomezeně
- <mods:extent>- neomezeně
- <mods:start>- výskyt právě jednou
- <mods:end>- výskyt právě jednou
- <rating>- výskyt právě jednou

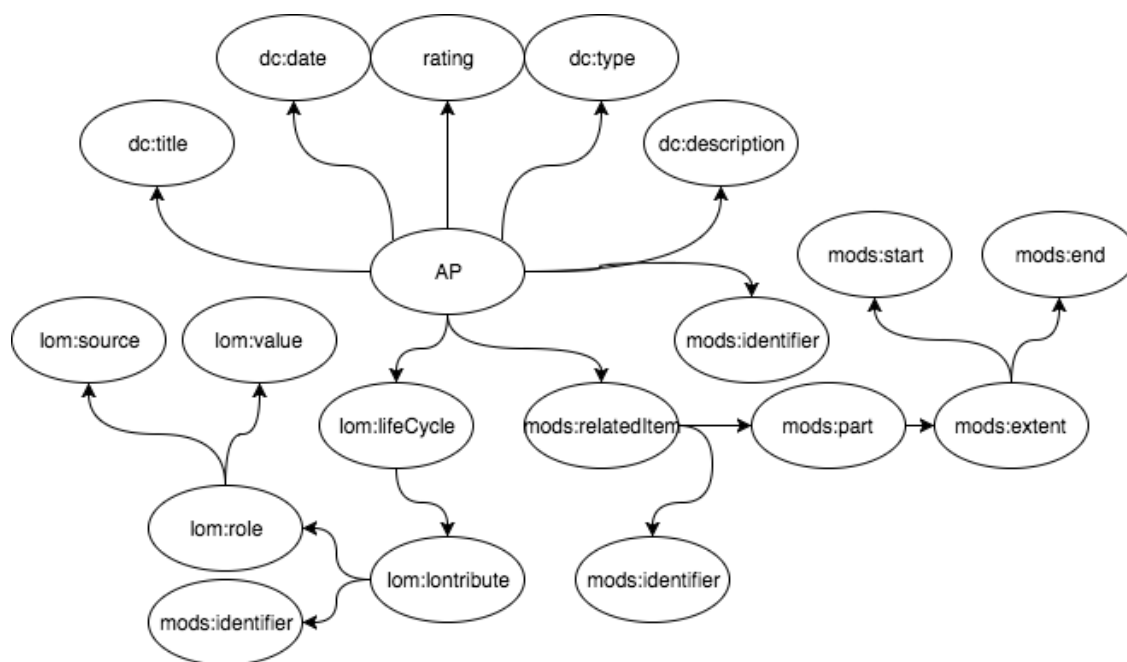
Fáze 5. Definice datových a výčtových typů

Upřesnění atributu <mods:identifier> *type* je dle modelu MODS datového typu string (text). Pro potřeby tohoto aplikačního profilu bude změněn na výčtový typ nabývacích hodnot množiny používaných identifikátorů digitálních objektů a autorů (viz oddíly 4.5.2, 4.5.3) a to: *DOI*, *ISBN*, *ISSN*, *researchID*, *scopusID*, *ORCID*. Elementu <dc:subject> byl definován datový typ „dcterms:AGROVOC“.

Ostatním elementům budou zachovány datové typy dle jejich definice v modelech LOM, MODS a DC.

Fáze 6. Definice závislostí mezi elementy

Závislosti mezi elementy jsou vyjádřeny v konceptuálním diagramu (viz [Obrázek 17](#)).



Obrázek 17 – schéma konceptuální datového modelu - zdroj: Autor.

Fáze 7. Publikování aplikačního profilu

Bylo vytvořeno XML schéma aplikačního profilu (viz příloha G), a tím ověřena navržená metodika MSOD. Dále byl vytvořen diagram konceptuálního datového modelu aplikačního profilu. V průběhu tvorby aplikačního profilu byla upravována navrhovaná metodika - použití metody pro zlepšování PDSA viz metodický postup této práce.

6.2 Tvorba záznamu digitálního artefaktu (objektu)

S použitím vytvořené XML šablony byl vytvořen XML záznam (viz příloha H), popisující konkrétní vědecký článek, a tím prakticky ověřena navržená metodika MSOD. Během tvorby XML záznamu byla průběžně upravována navrhovaná metodika, následně XML schéma - na základě použití metody pro zlepšování PDSA, viz metodický postup této práce.

Výsledný XML je čitelný a vhodný k využívání lidmi i stroji (Human-readable, Machine-readable).

6.3 Omezení pro použití navrhované metodiky

Metodika MSOD řeší problematiku vytvoření aplikačního profilu pouze pro metadatový popis digitálního artefaktu. V aplikacích typu sociální síť pro vědecké pracovníky se vyskytují profily dalších objektů (osoby, instituce), popis metadat těchto dalších objektů navrhovaná metodika neřeší.

Aplikace typu sociální síť pro vědecké pracovníky jsou převážně provozovány komerčními subjekty. Vůle pro zpřístupnění funkcí na export metadat pro uživatele může být limitována obavou z odchodu uživatelů ke konkurenčním aplikacím. Obecně lze konstatovat, že komerční provozovatelé upřednostňují získávání dat při jejich poskytování.

Vývoj v oblasti sociálních sítí je velmi dynamický, potřeba rychlých a kvalitních informací ve vědeckých komunitách bude vyžadovat další změny v metodice vzhledem k jejich vývoji.

Řada vydavatelů vědeckých časopisů není otevřených (Open Access Journals), metodika umožní sdílení metadat o digitálním artefaktu, ale přístup k plnému obsahu je omezen.

Navržená metodika vytváří aplikační model ve formátu schéma XML. Schéma je možno transformovat do formátu RDF nebo JSON schéma.

Pro vytvoření záznamu je třeba minimálně dvou elementů popisu, a to název digitálního artefaktu a jeho typ (text, vizualizace, video,...).

6.4 Zhodnocení a přínosy navrhované metodiky

V souladu se stavem poznání v dané oblasti vědecké činnosti byla navržena nová metodika dle zásad vědecké práce. Navržená metodika (MSOD) byla ověřena a ukázala se být přínosem v praxi. Metodika je originální a umožňuje výměnu obohacených dat mezi aplikacemi typu sociální síť pro vědecké pracovníky.

7 Závěry disertační práce

7.1 Splnění dílčích cílů disertační práce

Předkládaná disertační práce řeší aktuální problematiku v oblasti výměny vědeckých informací a přinesla na vytyčené cíle tyto odpovědi – výsledky:

Prozkoumat současnou úroveň webových aplikací pro podporu vědecké činnosti z hlediska metadatových formátů popisu digitálních artefaktů (objektů) a jejich sdílení pomocí webových služeb:

Pro efektivní práci s digitálními objekty je velmi důležité jejich obohacení o popisná data, tzv. *metadata*.

Při sdílení digitálních dokumentů je sdílen pouze jejich metadatový popis. Velmi rozšířeným a vhodným nástrojem pro sdílení metadat je *OAI-PMH*. *OAI-PMH* využívá některé principy webových služeb.

Vhodný datový formát pro metadatový zápis je XML. Dále by bylo možné použít JSON, ale transformace je jen technická záležitost. JSON nemá konečnou verzi schémat pro validaci formátu.

Pro další analýzy byly vybrány následující tři metadatové modely vhodné pro popis digitálních artefaktů:

- DC (Dublin Core),
- LOS (Learning object metadata)
- MODS (Metadata Object Description Schema, a bibliographic description schema).

Prozkoumat současnou úroveň webových aplikací pro podporu vědecké činnosti z hlediska otevřeného obsahu:

Významným trendem v současném světě je publikování informací pod licencemi, umožňující jejich rychlé a snadné šíření. V současné době patří mezi nejvhodnější licence pro šíření digitálních dokumentů soubor licencí, vyvinutý organizací *Creative Common*.

Prozkoumat současnou úroveň webových aplikací pro podporu vědecké činnosti z hlediska stavu a funkcí aplikací typu vědecká sociální síť:

Metadatové zápisy z *institucionálních repositářů* se načítají (sklízají) a shromažďují v *centrálních repositářích*. Repositáře poskytují vyhledávací služby nad metadatovými zápisy z *institucionálních repositářů*. Současný trend je vytvářet nad repositáři *sociální sítě pro vědecké pracovníky*, které rozšiřují primární vyhledávací funkci repositáře o další komunikační funkce. V rámci aplikací typu sociální sítě pro vědecké pracovníky jsou uživatelé k jednotlivým objektům generována *obohacující data*, která nejsou zatím již dále sdílena.

Sociální sítě a jejich data jsou vhodným doplňkem LMS – integrací sociálních sítí do LMS umožní prohloubení kvality sdílení znalostí a komunikace mezi uživateli a tvůrci materiálů. Dále umožní snadným způsobem přístup uživatelů k novým poznatkům daného oboru.

Při své činnosti vytváří sociální sítě mnoho metadat, kterými obohacují původní záznamy (obohacující data). Obohacující data byla na základě analýz klasifikována do dvou skupin, a to *vazební* a *další metadata*. Strukturu a funkci *dalších metadat* si vytváří každá sociální síť zvlášť. Příkladem těchto dat mohou být například komentáře, hodnocení (rating) apod. Mezi vazební metadata patří vazby:

- Digitální artefakt <-> Osoba (autor, spolutvůrci)
- Digitální artefakt <-> Digitální artefakt (citace, reference)

Analyzovat vybrané institucionální repositáře a způsoby poskytování dat a identifikaci subjektů v aplikacích typu sociální síť pro vědecké pracovníky (Social network service for scientists):

Celkem bylo analyzováno 49 vědeckých časopisů. Bylo zjištěno, že žádný z vybraných zkoumaných repositářů neposkytuje svá data pomocí veřejného API. Oproti tomu časopisy poskytují svá metadata pomocí

protokolu OAI-PMH knihovnám, např. Knihovně AV ČR, v. v. i. Všechny repositáře poskytují vyhledávací funkce, z čehož lze usuzovat, že publikované články popisují metadata a disponují nějakým aplikačním protokolem. Třicet šest z hodnocených repositářů poskytují obsah bez poplatku a jsou licencované některou z otevřených licencí, lze je tedy zařadit do kategorie otevřené časopisy (Open access journal).

Nejednotnost identifikace autora je stále problémem převážně v systémech, které publikují digitální artefakty. Identifikátory, které jsou čím dál více rozšiřovány a využívány jsou ORCID – identifikace autorů, DOI – identifikace digitálních objektů. ORCID se snaží spojit dva proprietární identifikátory researchID a Scopus Autor ID. Pro navrhovanou metodiku se doporučuje uvést všechny dostupné identifikátory, zkvalitní to výsledný záznam.

Zhodnotit problematiku a možnosti zápisu identifikace autorů, referencí a citací vybranými metadatovými modely:

Vybrané metadatové modely DC, LOM a MODS byly analyzovány a byly zjištěny následující závěry a doporučení pro navrhovaný metodický postup:

- DC je vhodný formát pro zápis základních metadat.
- Pro popis vazeb na osoby je vhodný standard LOM doplněný elementem identifier z modelu MODS.
- Pro popis vazeb na digitální artefakt je vhodný model MODS.
- Popis dalších obohacujících dat je nutno vytvořit vzhledem k jejich povaze.
- Schémata lze mezi sebou mapovat / transformovat.

Navrhnout postup pro zkvalitnění práce s daty vzniklými činností aplikací typu sociální síť pro vědecké pracovníky:

V této disertační práci byl navržen metodický postup MSOD (metodika sdílení obohacujících dat). Použitím této metodiky vznikne aplikační profil,

na základě kterého lze vytvořit konkrétní metadatové záznamy konkrétních digitálních artefaktů. Samotná metodika se skládá ze dvou etap. V první přípravné etapě jsou identifikována metadata digitálního artefaktu (objektu), druhá etapa se skládá ze sedmi fází (viz obrázek **Obrázek 18**), sloužících k vytvoření schématu samotného aplikačního profilu.

Ověřit správnost postupu vytvořením aplikačního profilu digitálního artefaktu (objektu):

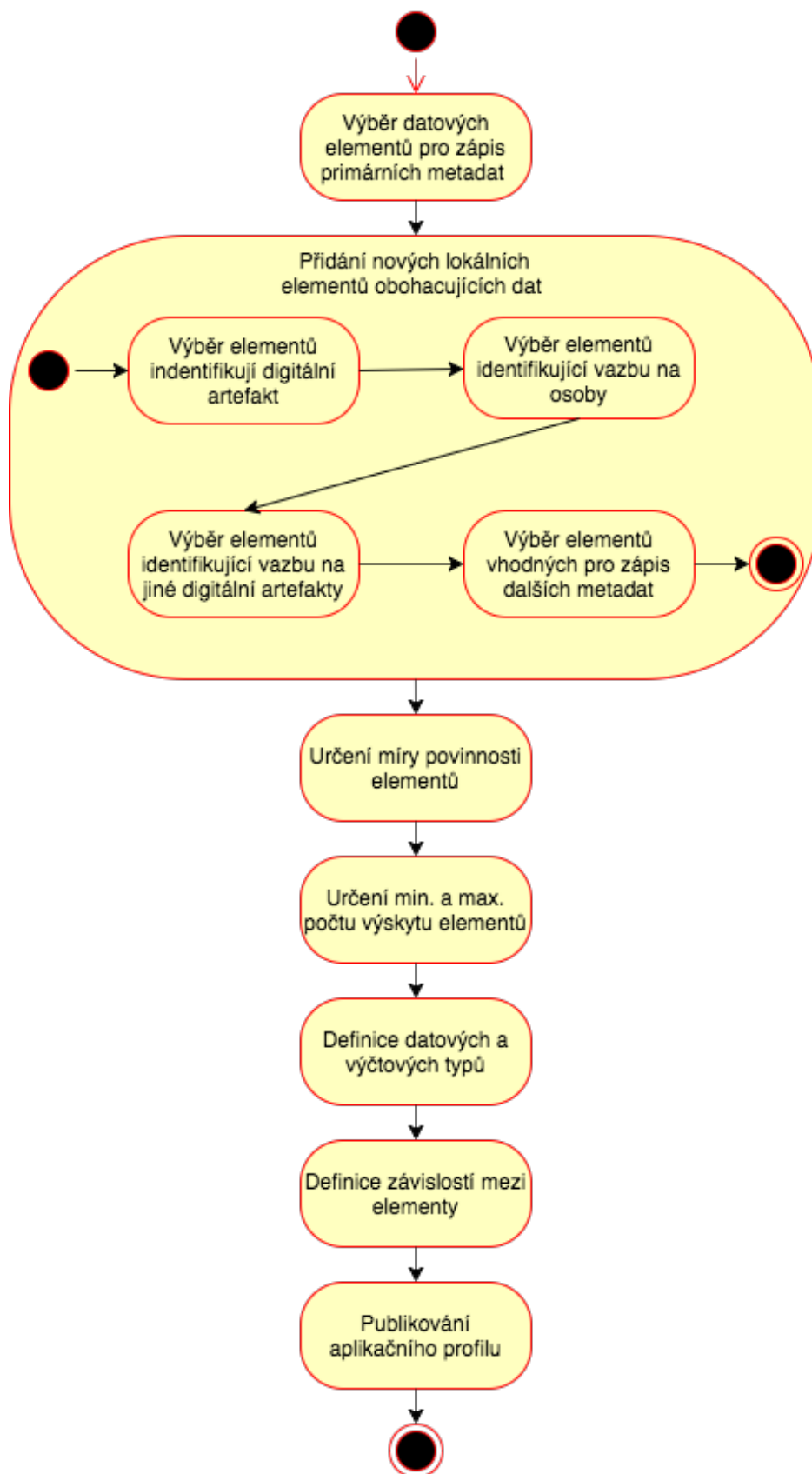
Na základě navrženého metodického postupu byl vytvořen aplikační profil pro sdílení digitálních artefaktů (objektů), sloužící k ověření správnosti metodického postupu. Výsledný aplikační profil je prezentován pomocí textového popisu a XSD XML schématu.

Na základě poznatků při tvorbě aplikačního profilu byl modifikován navržený metodický postup (dílčí cíl 4.).

Vytvořit záznam vybraného digitálního artefaktu na základě pravidel vytvořeného aplikačního profilu:

Dalším krokem ověření metodického postupu bylo vytvoření záznamu vybraného digitálního artefaktu. Záznam byl tvořen využitím pravidel aplikačního profilu ve formátu XML. Dále bylo ověřeno, že výsledný záznam je čitelný a vhodný k využívání lidmi i stroji (Human-readable, Machine-readable).

Na základě poznatků při tvorbě záznamu byl modifikován aplikační profil, respektive navržený metodický postup (dílčí cíl 4.).



Obrázek 18 – MSOD - diagram aktivit tvorby aplikačního profilu.

Zdroj: Autor

Formulovat závěry včetně zhodnocení navrženého postupu:

V souladu se stavem poznání v dané oblasti vědecké činnosti byla navržena nová metodika dle zásad vědecké práce. Navržená metodika (MSOD) byla ověřena a ukázala se přínosná pro praxi. Metodika je originální a umožňuje výměnu obohacených dat mezi aplikacemi typu sociální síť pro vědecké pracovníky. Dynamický rozvoj sociálních sítí a potřeba rychlých a kvalitních informací ve vědeckých komunitách bude vyžadovat další změny v metodice vzhledem k vývoji.

Tabulka 9 rekapituluje vztahy dílčích cílů ke kroku metodického postupu disertační práce a místa (oddílu), kde je daná problematika řešena.

Krok metodického postupu disertační práce	Dílčí cíl disertační práce	Oddíl disertační práce
1. krok	1.a	4.2
	1.b	4.3
	1.c	4.4
2. krok	2.	4.5
3. krok		
4. krok	3.	4.6
5. krok	4.	5
6. krok	5.	6.1
7. krok	6.	6.2
8. krok	7.	7

Tabulka 9 – Rekapitulace cílů vzhledem k oddílům. Zdroj: Autor.

7.2 Přínosy disertační práce pro vědu a praxi

Hypotéza disertační práce, zformulovaná na základě studia teoretických východisek, konstatuje, že dosud nebyl publikován metodický postup pro sdílení metadat aplikací typu sociální sít' pro vědecké pracovníky. Navržená metodika částečně doplňuje tuto mezeru.

Hlavním přínosem metodiky MSOD je zjednodušení zkvalitnění přenosu metadatových popisů mezi aplikacemi typu sociální sít' pro vědecké pracovníky.

7.3 Diskuse a náměty k dalšímu vědeckému výzkumu

Výzkumná oblast sdílení aplikací typu sociální sít' nabízí několik dalších oblastí ke zkoumání.

Další možnosti výzkumu se nacházejí v rozšíření a úpravách navržené metodiky MSOD. Metodika řeší problematiku vytvoření aplikačního profilu pouze pro sdílení metadat digitálního artefaktu. V aplikacích typu sociální sít' pro vědecké pracovníky se vyskytují profily dalších objektů (osoby, instituce). Začlenění popisu těchto dalších objektů do aplikačního profilu by umožnilo uživateli exportovat veškeré aktivity a data vytvořená v aplikaci. Takto standardizované záznamy by byly nezávislé na platformě sociální sítě a bylo by je možno využít k snadnému přechodu mezi různými sociálními sítěmi nebo k archivaci záznamů pro další využití.

Navržená metodika MSOD byla konstruována pro využití v oblasti sociálních sítí pro vědecké pracovníky. Dále by bylo možné navrženou metodiku zobecnit a upravit pro zkvalitnění přenosu metadat ve všech typech sociálních sítí.

8 Seznam literatury

- ABRAHAMSE, B., *Getting Started with MarcEdit*, New England Library Association Fall 2013 Conference 2013, [online]. [cit. 2017-02-09]. Dostupné z <https://netsl.files.wordpress.com/2013/04/getting-started-with-marcedit-tuesday-nela-2013.pdf>
- ALIJANI, A.S., JOWKAR, A. *Dublin Core Metadata Element Set usage in national libraries' web sites*: Electronic Library. 2009. roč. 27, č.3, s. 441-447.
- ARMS, W. Y. *Digital Libraries*. Cambridge: MIT Press, 2000. ISBN 0-262-01880-8.
- BARLEY, P., *Extension of programming language concepts*. Philadelphia: University City Science Center, 1968.
- BBC Creative Archive pilot*. BBC. [online]. [cit. 2016-10-09]. Dostupné z: <http://www.bbc.co.uk/creativearchive/>.
- BARTOŠEK, M. *Digitální knihovny – teorie a praxe*, Národní knihovna ČR 2004, roč. 15, č. 4, s. 233-254. ISSN 1214-0678.
- BARTOŠEK, M., *Vyhledávání v Internetu a DUBLIN CORE*. Zpravodaj ÚVT MU. ISSN 1212-0901, 1999, roč. IX, č. 4, s. 1-4.
- BARVAFLIA, G., MURZILLY, S., CUDINI, S. Definition of REST web services with JSON schema. *Softw. Pract. Exper.*, 2017 č.47: s.907–920. doi: 10.1002/spe.2466.
- BERAN, Jiří. *Metadata*. [online]. [cit. 2016-12-10]. Dostupné z: <http://kisk.phil.muni.cz/wiki/Metadata>
- BERNERS-LEE, T. *Linked Data*. [online] 2006-07-27, aktualizováno 2009-06-18 [cit. 2016- 10-20]. Dostupné z: <http://www.w3.org/DesignIssues/LinkedData.html>.
- BOHÁČEK, M. *Licenční smlouvy v oblasti autorských práv*. Praha : Metropolitní univerzita Praha, 2008. s. 9-35. ISBN: 978-80-86855-29-5.
- BÖHM, C. A kol. : *Integrating Open Government Data for transparency*. Publikováno v *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion*. Lyon: 2012. s. 321 – 324. ISBN 978- 145031230-1.
- Budapest Open Access Initiative*. [online]. [cit. 2016-12-10]. Dostupné z: <http://www.budapestopenaccessinitiative.org/>
- BRAND, A., DALY, F., MEYERS, B. *Metadata demystified*. Bethesda: 2003 NISO Press. ISBN 1-880124-59-9.
- BRATKOVÁ, E. *Metadata jako nový nástroj pro komunikaci webovských informačních zdrojů*, Národní knihovna: 1999, roč. 10., č.4., s. 178-195 ISSN 0862-7487
- BRATKOVÁ, E., MACH, J. *Metadatový soubor pro elektronické vysokoškolské kvalifikační práce v ČR*. Asociace knihoven vysokých škol ČR: 2008.
- BRETHERTON P., F., SINGLEY. P. T. Metadata: a user's view. In *Proceedings of the 7th international conference on Scientific and Statistical Database*

- Management* (SSDBM'1994), James C. French and Hans Hinterberger (Eds.). IEEE Computer Society, Washington, DC, USA, 1994, s. 166-174.
- CARPENTER, L. *Main Technical Ideas of OAI-PMH*. [online]. [cit. 2016-10-09]. Dostupné z: <http://www.oaforum.org/tutorial/english/page3.htm>
- Codes for the representation of names of languages*, ISO 639-3:2007, 2007.
- DE ALMEIDA, D.B.F.C a GUERRA E.M., *Evolution of XSD Documents and Their Variability During Project Life Cycle: A Preliminary Study 2016* s. 392 . DOI: 10.1007/978-3-319-42089-9_28.
- DESCARTES, R. *Rozprava o metodě*. Praha: Svoboda, 1992. 67 s. ISBN 80-205-0216-5.
- CHANG, W. *Embedding MPEG-7 metadata within a media file format*, Proceedings of SPIE - The International Society for Optical Engineering: 2005, s. 1-12.
- CEJPEK, J. *Dějiny knihoven a knihovnictví*. 2. dopl. vyd. V Praze: Karolinum, 2002, 247 s., ISBN 8024603233.
- CELBOVÁ, L. *Dublin Core Metadata Element Set: Stručný návod k použití s příklady*. Praha: Národní knihovna ČR, 2003. 14 s.
- Common Format and MIME Type for Comma-Separated Values (CSV) Files RFC 4180*, 2005.
- Creative Commons Česká republika* [online]. [cit. 2017-03-03]. Otázky pro autory. Dostupné z: <http://www.creativecommons.cz/483/otazky-pro-autory/>.
- CUBR, L., HAVLOVÁ, J.. *Digitální repozitář. KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) Praha : Národní knihovna ČR, 2003-* [online].[cit. 2016-12-10]. Dostupný z: http://aleph.nkp.cz/F/?func=direct&doc_number=000014292&local_base=KTD.
- DCMI Type Vocabulary*, DCMI, 2010, [online].[cit. 2016-12-10]. Dostupné z <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>
- DEMING, E.W., *PDSA Cycle*, The W. Edwards Deming Institute , [online].[cit. 2016-02-01]. Dostupné z <https://deming.org/management-system/pdsacycle>
- Digitální knihovna ČVUT*, Ústřední knihovna ČVUT., [online].,[cit. 2016-12-10]. Dostupný z: <http://knihovna.cvut.cz/informacni-zdroje/institucionalni-repozitar/>
- DONOHUE, T. *Documentation*, DSpace , 2015, [online].,[cit. 2016-12-10]. Dostupné z: <https://wiki.duraspace.org/display/DSPACE/Documentation>
- DOUCEK, P., MARYŠKA, M. a NEDOMOVÁ, L. 2013. *Informační management v informační společnosti*. Praha : Edition (c), 2013. ISBN 978-80-7431-097-3.
- DRUCKER, P. F. *Nové reality*. Praha: Management Press, 1995. 1. vyd. 244 s. ISBN 80-85603-85-3.
- Dublin Core : Czech, 2006, [online].,[cit. 2016-12-10]. Dostupný z: http://www.ics.muni.cz/dublin_core/index.html.
- Dublin Core Metadata Element Set*, ANSI/NISO Z39.85-2007. ISSN 1041-5635, 2007.

- Documnetation*, ePrint, [online],[cit. 2016-12-10]. Dostupný z: http://wiki.eprints.org/w/Main_Page
- EHLERS, U. , M. PAWLOWSKI, J. *Handbook on quality and standardisation in E-learning*. Berlin: Springer, 2006. ISBN 978-3-540-32787-5.
- EuroVoc, *EuroVoc, mnohojazyčný tezaurus Evropské unie* , [online],[cit. 2016-12-10]. Dostupný z: <http://eurovoc.europa.eu/drupal/?q=cs>
- Extensible Markup Language (XML)*, W3C, [online],[cit. 2016-12-10]. Dostupný z: <http://www.w3.org/XML/>
- FAO AgriDrupal*, FAO , [online],[cit. 2016-12-10]. Dostupný z: <http://aims.fao.org/vest-registry/tools/agridrupal>
- FAO AIMS* FAO , [online],[cit. 2016-12-10]. Dostupný z: <http://aims.fao.org/>
- FEGEN, N. *What is the OAI Protocol for Metadata Harvesting*. [online]. [cit. 2016-09-09]. Dostupné z: http://wiki.cetis.ac.uk/What_is_the_OAI_Protocol_for_Metadata_Harvesting
- FERNÁNDEZ, J. D. a kol. *Binary RDF representation for publication and exchange (HDT)*. Publikováno v *Journal of Web Semantic*, č. 19, s 22 – 41. Elsevier, 2013. ISSN 1570-8268.
- Fond Otakara Motejla. *Co jsou otevřená data*. [online] [cit. 2016-10-20]. Dostupné z <http://www.otevrenadata.cz/otevrena-data/co-jsou-otevrena-data/>.
- IMS Global Learning Consortium. *IMS Application Profile Guidelines Overview*, . [online] [cit. 2017-02-03]. Dostupné z http://www.imsglobal.org/ap/apv1p0/imsap_oviewv1p0.html, 2005.
- Internet Media Type registration, consistency of use*, W3C, 2002, [online],[cit. 2016-12-10]. Dostupné z: <http://www.w3.org/2001/tag/2002/0129-mime>.
- ISO/IEC JTC1/SC7 3941. FCD 24765 - Systems and software engineering vocabulary: 2008 [online] [cit. 2016-10-20]. Dostupné z http://jtc1-sc7.logti.etsmtl.ca/N4101%20-%20N4150/07N4146%20WG22_MeetingMinutes_Berlin_052008.pdf
- GAROFALO, D.A. *E-Book Cataloging (NCSU Libraries)*. Technical Services Quarterly, 2015. roč. 32. č.1, s. 96.
- GRUBER, L. Creative Commons. LarPard. 2011 [online] [cit. 2016-10-20]. Dostupné z <http://www.larpard.cz/cc>
- HENDL, J. 2008. *Kvalitativní výzkum: základní teorie, metody a aplikace*. 2., aktualizované vydání Praha: Portál, 2008, 407 s. ISBN 978-80-7367-485-4.
- HEERY, R., PATEL M. *Application Profiles: Mixing and Matching Metadata Schemas*, 2000 ARIADNE, vol. 25, John Wiley & Sons
- High-level framework for network-based resource sharing*, RFC 707, 1976.
- HODGE, G., TEMPLETON, C. ALLEN, R. *A metadata element set for project documentation*. Science and Technology Libraries: 2005, roč. 25 č.4, s. 5-23.

HOFMAN, W. RAJAGOPAL, M. *A technical framework for data sharing*. Publikováno v Journal of Theoretical and Applied Electronic Commerce Research, ročník 9, č. 3, s 45 – 58. University de Talca, 2014, Talca. ISSN 0718-1876.

HONG, K., *Web Technologies – open APIS, SOAP, and REST 2015.*, 2005. [online]. [cit. 2016-12-10]. Dostupné z: http://www.bogotobogo.com/WebTechnologies/OpenAPI_RESTful.php

HÜLLEN, W., *A History of Roget's Thesaurus: Origins, Development, and Design*. 2004 : Oxford University Press. ISBN 978-0-19-928199-2

LLOYD, S., ROGET, P., M. *Roget's thesaurus of English words and phrases*. Harlow, Essex: Longman, 1982, 1247 p. ISBN 0582555515x.

KAHN, R., WILENSKY, R. A Framework for Distributed Digital Object Services. Technical Report hdl:cnri.dlib/tn95-01. CNRI, 1995. , [online]., [cit. 2016-12-10]. Dostupný z: <http://www.cnri.reston.va.us/home/cstr/arch/k-w.html>.

KENNEDY, B., MUSCINO, C. *HTML & XHTML: The Definitive Guide*, 6th Edition. O'Reilly, 2006. ISBN: 0596527322.

KINTISCH, E., *Is ResearchGate Facebook for science?, 2014.* , [online]., [cit. 2016-12-10]. Dostupný z: http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2014_08_25/caredit.a1400214

KLUG, H., KMOCH A. *Operationalizing environmental indicators for real time multi-purpose decision making and action support*. Článek bude publikován v Ecological Modelling. Elsevier, 2014. ISSN 0303-3800

KOZAKI, K. , YAMAGATA, Y., IMAI, T. , OHE, K. – *Mizoguchi, R. Publishing a disease ontologies as linked data*. Publikováno v *Lecture Notes in Computer Science*, č. 8726, s 110 – 128. Springer Verlag, 2014. ISSN 0302-9743

KUBA, M. *Web Services*. DATAKON 2006: 2006, s. 336 ISBN 80-210-4102-1.

KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. [cit. 2008-11-28]. Praha: Národní knihovna České republiky, 2003. Dostupné z URL: <<http://sigma.nkp.cz/cze/ktd>>

MARBI.: *MARC 21 Formats*, Library of Congress, , [online]., [cit. 2016-12-10]. 1996, Dostupné z: <http://www.loc.gov/marc/96principl.htm>.

Metadata Standard, Collective Access, 2004, [online]., [cit. 2016-12-10]. Dostupné z: http://docs.collectiveaccess.org/wiki/Metadata_Standards

MLÝNKOVÁ, I., POKORNÝ, J. *Technologie XML: Principy a aplikace v praxi*. Praha : Grada, 2008. ISBN 978-80-247-2725-7.

MODS User Guidelines Version 3, Library of Congress, 2009, [online]., [cit. 2016-12-10]. Dostupné z: <http://www.loc.gov/standards/mods/v3/mods-userguide.html>

MØLLER, A., SCHWARTZBACH, M., *An introduction to XML and Web technologies: The manual*. Addison-Wesley, 2006. str. 542. ISBN 978-03-2126-9669.

MÜLLER, T. H., ADELHARD, K. *A web-based central diagnostic data repository*. Paper presented at the Studies in Health Technology and Informatics, (2002), č. 90, s.246-250.

Nature, Credit where credit is due. Naturel 2009, 462(7275), 825-825 [cit. 2017-05-04]. DOI: 10.1038/462825a. ISSN 0028-0836. Dostupné z: <http://www.nature.com/doifinder/10.1038/462825a>

NILSSON M., MILES A.J., JOHANSON P., ENOKSSON F. *Formalizing Dublin Core Application Profiles – Description Set Profiles and Graph Constraints*. 2009 DOI: 10.1007/978-0-387-77745-0_10

NFS Version 3 Protocol Specification, RFC 1813, 1995.

NONTHAKARN C., WUWONGSE V., An application profile for research collaboration and information management, 2015m roč.. 49 č.: 3, str..242-265, DOI: 10.1108/PROG-02-2014-0007

Otevřený přístup (open access) [online]. [cit. 2016-09-09]. Dostupné z: <http://www.openaccess.cz/>.

Open archival information systém, ISO 14721:2003

Open Archive Initiative. The Open Archives Initiative Protocol for Metadata Harvesting. [online]. [cit. 2016-09-09]. Dostupné z: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

Open Definition, *GNU Free Documentation License (GFDL)* [online]. [cit. 2017-02-03]. Dostupné z: <http://opendefinition.org/licenses/gfdl/>.

PLANKOVÁ, J. *Technická řešení pro otevřený přístup*. Ikaros: [online]. 2008, roč. 12, č. 2 ISSN 1212-5075.

PMH Tools. [online]. [cit. 2016-10-09]. Dostupné z: <http://www.openarchives.org/pmh/tools/tools.php>

Representation of dates and times, ISO 8601:1988, 1988.

RAO, A. a další. 1996. Total Quality Management: A Cross Functional Perspective. místo neznámé: John Wiley & Sons, 1996. ISBN 0-471-10804-9.

SGOUROPOULOU, C., a kol. *Specification of metadata profiles and mappings to existing technology* (Part B: The VOA3R AgRes AP Metadata Terms). Deliverable number D 3.5 of Virtual Open Access Agriculture & Aquaculture Repository. Technological Educational Institute of Athens, 2011.

SCHIERNEIR, Q. *Initiative aims to break science's citation paywall*. Nature. 2017 ISSN 1476-4687. doi:10.1038/nature.2017.21800.

SCHWARZ, J., SCHWARZ, A. *Aktualizace tezauru EUROVOC a poloautomatická reindexace*, Ikaros: [online]. 1999, roč. 3, č.1. ISSN 1212-5075.

SIOLVA, T. – WUWONSE, V. – SHARMA, H. N. *Disaster mitigation and preparedness using linked open data*. Publikováno v Journal of Ambient Intelligence and Humanized Computing, ročník 4, č. 5, s 591 – 602. Springer Verlag, 2013. ISSN 1868-5137.

SMITH, N. VAN COILLIE M., DUVAL E., Guidelines and Support for Building Application Profiles in E-Learning, Proc. CEN/ISSS WS/LT Learning Technologies Workshop CWA, pp. 1-26, 2016.

SOUNDARARAJAN, E. a kol. *Semantic digital library - Migration of Dublin core to RDF*, Proceedings of the 2010 International Conference on Signal and Image Processing, ICSIP 2010, 2010, s. 250-254.

SPAMPINATO, D. ZAMGARA, I. *Classical antiquity and semantic content management on linked open data*. Publikováno v 1st International Workshop on Collaborative Annotations in Shared Environment: Metadata, Vocabularies and Techniques in the Digital Humanities. Florence, 2013. ISBN 978-145032199-0.

SRIPARASA, S. *JavaScript and JSON Essentials: Successfully Build Advanced JSON-fueled Web Applications with This Practical, Hands-on Guide*. Birmingham, UK: Packt, 2013. ISBN: 9781783286034.

Standard Generalized Markup Language (SGML), ISO 8879:1986, 1986.

STOČES, M., MASNER, J., JAROLÍMEK, J. Mitigation of Social Exclusion in Regions and Rural Areas – E-learning with Focus on Content Creation and Evaluation. AGRIS on-line Papers in Economics and Informatics, 2015, roč. 7, č. 4, s. 143-150. ISSN: 1804-1930.

ŠIMEK, P. a kol. *Using metadata formats and AGROVOC thesaurus for data description in the agrarian sector*. Plant, Soil and Environment, 2013, roč. 59, č. 8, s. 378-384. ISSN: 1214-1178.

ŠIMEK, P. a kol. *Using of Automatic Metadata Providing*. AGRIS on-line Papers in Economics and Informatics, 2013, roč. V, č. 4, s. 189-197. ISSN: 1804-1930.

ŠIMEK, P. a kol. *Using Metadata Description for Agriculture and Aquaculture Papers*. AGRIS on-line Papers in Economics and Informatics, 2012, roč. IV, č. 4, s. 79-90. ISSN: 1804-1930.

ŠIROKÝ, J. 2011. *Tvoříme a publikujeme odborné texty*. Vyd. 1. Brno: Computer Press, 2011, 208 s. ISBN 978-80-251-3510-5.

ŠVÁSTOVÁ, P. *MODS a další metadatová schémata v oblasti digitalizace dokumentů*. Knihovny současnosti 2009: 2009, roč. 17, s. 127 - 132 - ISBN 978-80-86249-54-4.

TAHERI, S.M., HARIRI, N., *A comparative study on the indexing and ranking of the content objects including the MARCXML and Dublin Core's metadata elements by general search engines*. Electronic Library, 2012, roč. 30, č.4, s. 480-491.

TAYLOR, A., G., JOUDREY, D.,N. *The organization of information. 3rd ed. Westport, Conn.:* Libraries Unlimited, 2009, roč. 16, s. 512 p. ISBN 159158700x.

Tezaurus, Easylibrary, 2010. . [online]. [cit. 2015-04-11]. Dostupné z: <http://easylibrary.webnode.cz/news/tezaurus/>.

The Dublin Core metadata element set, ISO 15836:2009. 2009.

The Dublin Core Metadata Initiative (DCMI), [online]. [cit. 2016-12-10]. Dostupné z: <http://dublincore.org/>.

The eXtensible Catalog (XC) Drupal Toolkit, Drupal, 2015. , [online]. [cit. 2016-12-10]. Dostupné z: <https://www.drupal.org/project/xc>.

The Singapore Framework for Dublin Core Application, [online]. [cit. 2016-12-10]. Dostupné z: [Profileshttp://dublincore.org/documents/singapore-framework/](http://dublincore.org/documents/singapore-framework/)

The Roots of SGML, A Personal Recollection, [online]. [cit. 2016-12-10]. Dostupné z: <http://www.sgmlsource.com/history/roots.htm>.

VAN DE SOMPEL, H., HOCHSTENBACH, P. *Reference Linking in a Hybrid Library Environment. Part 3, Generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment*. D-Lib Magazine: 1999, roč. 5, č. 10. [online]. [cit. 2016-12-10]. Dostupný z:

http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html.

VAŇKOVÁ, J. *Teorie informace v gymnaziálním kurzu informatiky*. Metodický portál: Články, 2011, ISSN 1802-4785. [online]. [cit. 2015-04-11]. Dostupné z: <http://clanky.rvp.cz/clanek/c/g/14121/TEORIE-INFORMACE-V-GYMNAZIALNIM-KURZU-INFORMATIKY.html>.

VODÁČEK, L., ROSICKÝ, A. *Informační management - pojetí, poslání a aplikace*. Praha: Management Press 1997. ISBN 80-85943-35-2.

WEBBER, J., PARASTATIDIS, S., ROBINSON, I. *REST in Practice*. Farnham: O'Reilly, 2010. ISBN:978-0-596-80582.

W3C. Resource Description Framework (RDF). [online] [cit. 2017-03-02] Dostupný z: <http://www.w3.org/RDF/>.

Zemědělská a potravinářská knihovna. *Tezaury*. [online] [cit. 2016-09-09]. Dostupné z <http://www.nzpk.cz/tezaury/>.

9 Seznam příloh

A.	SEZNAM OBRÁZKŮ.....	105
B.	SEZNAM TABULEK	106
C.	SEZNAM GRAFŮ.....	107
D.	SEZNAM POŽITÝCH ZKRATEK.....	108
E.	DEFINICE POUŽITÝCH POJMŮ	110
F.	ANALÝZA INSTITUCIONÁLNÍCH REPOSITÁŘŮ.....	111
G.	SCHÉMA APLIKAČNÍHO PROFILU.....	115
H.	XML ZÁZNAM	117
I.	PUBLIKAČNÍ ČINNOST AUTORA	119

A. Seznam obrázků

Obrázek 1 – Metoda PDSA. Zdroj: Autor.....	12
Obrázek 2 – PDLA ve vztahu s kroky postupu vypracování disertační práce. Zdroj: Autor.....	13
Obrázek 3 – Závislosti mezi „metodologií vědy“, „metodou vědy“ a „metodikou vědecké práce“. Zdroj: Autor dle (Ochrana, 2010).....	15
Obrázek 4 – Vztah pojmů data, informace, znalost. Zdroj: Autor	17
Obrázek 5 – Ukázka záznamu ve formátu MARC (Abrahamse, 2013).....	26
Obrázek 6 – Součásti profilu Dublin Core Application (Singapore Framework, 2017)	31
Obrázek 7 Ukázka výpisu záznamů "hovězí maso" z tezauru AGROVOC. Zdroj: FAO.	35
Obrázek 8. – Ukázka popisu dokumentu RDF (W3C, 2017).....	42
Obrázek 9 – princip OAI-PMH. Zdroj: Autor.....	48
Obrázek 10 – OAI-PMH pojmy. Zdroj:(Carpenter, 2013).	49
Obrázek 11 – příklad komunikace mezi centrálním a lokálním repositářem. Zdroj: (Carpenter, 2013)	50
Obrázek 12 – Referenční model OAIS. Zdroj: (Stoklasová, 2006).	59
Obrázek 13 – vztah institucionálních repositářů a oborově zaměřených soc. sítí. Zdroj: Autor.	64
Obrázek 14 – ORCID - výpis ID různých systémů. Zdroj: Autor.	69
Obrázek 15 – MSOD - diagram aktivit tvorby aplikačního profilu. Zdroj: Autor.....	82
Obrázek 16 – přenos metadat mezi aplikacemi. Zdroj: Autor.....	83
Obrázek 17 – schéma konceptuální datového modelu - zdroj: Autor.....	88
Obrázek 18 – MSOD - diagram aktivit tvorby aplikačního profilu. Zdroj: Autor ...	94

B. Seznam tabulek

Tabulka 1 – Vztah PDSA k metodě sedmi kroků (Rao, a další, 1996)	14
Tabulka 2 – DC - základní prvky. Zdroj: (Dublin Core : Czech, 2006).....	20
Tabulka 3 MARC – formáty. Zdroj: (Taylor, 2009).....	26
Tabulka 4 – Hlavní prvky schématu MODS. Zdroj: (MODS, 2009).....	27
Tabulka 5 – Přehled elementů VOA3R AP 4 zdroj (Sgouropoulou, 2011).....	32
Tabulka 6 – Analýza aplikací pro podporu vědy. Zdroj: Autor.	70
Tabulka 7 - mapování mezi DC a LOM.....	75
Tabulka 8 - mapování mezi DC a MODS.....	76
Tabulka 9 – Rekapitulace cílů vzhledem k oddílům. Zdroj: Autor.....	95
Tabulka 10 – Analýza institucionálních repositářů (Zdroj: Autor.).....	114

C. Seznam grafů

Graf 1 – Poměr API a technologií. Zdroj: (Hong, 2015).....	46
Graf 2 – REST a SOAP api 2006 a2016 Zdroj: https://www.programmableweb.com	46

D. Seznam použitých zkratk

AIMS	<i>Agriculture Information Management Standards</i>	<i>standardy pro zemědělský informační management</i>
ANCI/NISO	<i>American National Standards Institute / National Information Standards Organization</i>	
AP	<i>Application profile</i>	<i>aplikační profil</i>
API	<i>Application Programming Interface</i>	
CC	<i>Creative Commons</i>	<i>tvůrčí společenství</i>
COAR	<i>Confederation of Open Access Repositories</i>	
DC	<i>Dublin Core</i>	
DCMI	<i>Dublin Core Metadata Initiative</i>	
DSP	<i>Description Set Profile</i>	
DOI	<i>Digital Object Identifier</i>	
GNU	<i>GNU's Not Unix!</i>	
GNUF	<i>GNU Free Documentation License</i>	
HTML	<i>HyperText Markup Language</i>	
IEEE	<i>Institute of Electrical and Electronics Engineers</i>	
ISBN	<i>International Standard Book Number</i>	<i>mezinárodní standard číslování knih</i>
ISSN	<i>International Standard Serial Number</i>	<i>mezinárodní standard číslování periodické publikace</i>
ISO	<i>International Organization for Standardization</i>	<i>mezinárodní organizace pro normalizaci</i>
TCP / IP	<i>Transmission Control Protocol/Internet Protocol</i>	
JSON	<i>JavaScript Object Notation</i>	
LOM	<i>Learning Object Metadata</i>	<i>metadata učebních objektů</i>
MARC	<i>Machine-Readable Cataloging</i>	
MODS	<i>Metadata Object Description Schema</i>	<i>schema pro popis objektů</i>

MSOD	<i>Enriched Data Sharing Methodology</i>	<i>metodika sdílení obohacených dat</i>
OAI-PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>	
ORCID	<i>Open Research and Contributor ID</i>	
PDSA	<i>Plan-Do-Study-Act</i>	<i>plánuj-udělej-studuj-uskutečni</i>
RDF	<i>Resource Description Framework</i>	<i>system popisů zdrojů</i>
SOAP	<i>Simple Object Access Protocol</i>	
ST	<i>Singapore Framework</i>	
TEI	<i>Text Encoding Initiative</i>	
UML	<i>Unified Modeling Language</i>	
VOA3R	<i>Virtual Open Access Agriculture & Aquaculture Repository Project</i>	
WoS	<i>Web of Science</i>	
XML	<i>Extensible Markup Language</i>	<i>rozšiřitelný značkový jazyk</i>
XSD	<i>XML Schema Definition</i>	<i>iniciativa pro definici XML schématu</i>
WSDL	<i>Web Services Description Language</i>	
W3C	<i>World Wide Web Consortium</i>	

E. Definice použitých pojmů

Aplikace typu sociální síť pro vědecké pracovníky

Jedná se o aplikace, které slouží k podpoře vědecké činnosti, pracuje s metadaty digitálních artefaktů a při své činnosti metadata rozšiřuje (obohacuje) o další metadata. Tyto aplikace lze rozdělit do následujících skupin:

- Sociální sítě pro vědecké pracovníky
- Systémy pro zprávu referencí
- Vyhledavače odborných vědeckých prací

Digitální artefakt

Jedná se o digitální formu lidského výtvoru. Digitální artefakty mohou být reprezentovány především ve formě textů, vizualizací nebo zvuků nebo jejich kombinací. Též lze použít pojem digitální objekt.

Obohacující data

Jedná se o primární elementy digitálních artefaktů a data, která je obohatila činnostmi aplikací sociálních sítí pro vědecké pracovníky.

Primární elementy

Jedná se o elementy metadat, která vznikla při publikaci digitálního artefaktu. Jedná se například o název díla, autory, licence publikování apod.

F. Analýza institucionálních repositářů

Název časopisu	Impakt faktor	Open Access	Veřejné API	Zaměření
PRESLIA	3,000	ano	ne	botanika
Moravian Geographical Reports	2,149	ano	ne	geografie
BIOLOGIA PLANTARUM	1,551	ne	ne	botanika
FOLIA MICROBIOLOGICA	1,521	ne	ne	mikrobiologie
PHOTOSYNTHETICA	1,507	ne	ne	biologie
PHYSIOLOGICAL RESEARCH	1,461	ano	ne	fyzilogie
Journal of Applied Biomedicine	1,433	ne	ne	biomedicína
Fottea	1,350	ano	ne	psychologie
PLANT SOIL AND ENVIRONMENT	1,225	ano	ne	rostlinná výr.
BULLETIN OF GEOSCIENCES	1,175	ano	ne	geologie
EUROPEAN JOURNAL OF ENTOMOLOGY	1,167	ano	ne	entomologie
E & M Ekonomie a Management	1,163	ano	ne	Ekonomika a management
FOLIA PARASITOLOGICA	1,082	ano	ne	parazitologie
FOLIA GEOBOTANICA	1,017	ne	ne	geobotanika
Radioengineering	0.945	ano	ne	elektronika
FOLIA BIOLOGICA	0.939	ano	ne	biologie

Soil and Water Research	0.934	ano	ne	půda a země
BIOMEDICAL PAPERS- OLOMOUC	0.894	ano	ne	biomedicína
Agricultural Economics	0.789	ano	ne	Zemědělská ekonomika
CZECH JOURNAL OF FOOD SCIENCES	0.787	ano	ne	potraviny
STUDIA GEOPHYSICA ET GEODAETICA	0.764	ne	ne	geofyzika
PLANT PROTECTION SCIENCE	0.742	ano	ne	ochrana rostlin
CZECH JOURNAL OF ANIMAL SCIENCE	0.741	ano	ne	zoologie
FOLIA ZOOLOGICA	0.739	ano	ne	zoologie
Prague Economic Papers	0.710	ano	ne	ekonomie
Acta Geodynamica et Geomaterialia	0.699	ano	ne	geomechanika
Central European Journal of Public Health	0.682	ano	ne	medicína
ACTA ENTOMOLOGICA MUSEI NATIONALIS PRAGAE	0.632	ano	ne	entomologie
Slovo a Slovesnost	0.625	ne	ne	lingvistika
Applications of Mathematics	0.618	ne	ne	matematika
Journal of Geosciences	0.609	ano	ne	geologie

FINANCE A ÚVĚR	0.604	ano	ne	ekonomie
POLITICKA EKONOMIE	0.589	ano	ne	ekonomie
Geografie	0.580	ano	ne	geografie
HORTICULTURAL SCIENCE	0.566	ano	ne	zahradnictví
Acta Chirurgiae Orthopaedicae et Traumatologiae Cechoslovaca	0.560	ne	ne	traumatologie
Czech Journal of Genetics and Plant Breeding	0.532	ano	ne	rostlinná výroba
EPIDEMIOLOGIE MIKROBIOLOGIE IMUNOLOGIE	0.500	ne	ne	imunologie
VETERINARNI MEDICINA	0.489	ano	ne	veterinární medicína
CERAMICS-SILIKATY	0.439	ano	ne	keramika, sklo
ACTA VETERINARIA BRNO	0.415	ano	ne	veterinární medicína
Neural Network World	0.394	ano	ne	neuronové sítě
CHEMICKE LISTY	0.387	ano	ne	chemie
KYBERNETIKA	0.379	ano	ne	kybernetika
CESKA A SLOVENSKA NEUROLOGIE A NEUROCHIRURGIE	0.368	ano	ne	neurologie
CZECHOSLOVAK MATHEMATICAL JOURNAL	0.364	ano	ne	matematika

CESKOSLOVENSKA PSYCHOLOGIE	0.242	ne	ne	psychologie
LISTY CUKROVARNICKE A REPARSKÉ	0.218	ano	ne	rostlinná výroba - řepa
SOCIOLOGICKY CASOPIS-CZECH SOCIOLOGICAL REVIEW	0.143	ano	ne	sociologie

Tabulka 10 – Analýza institucionálních repositářů (Zdroj: Autor.)

G. Schéma aplikačního profilu

Pro přehlednost je kód rozčleněn na několik částí a zkrácen.

Schéma primárních dat:

```
<xs:element name="title" type="xs:string" minOccurs="1"
maxOccurs="1"/>
<xs:element name="type" type="xs:string" minOccurs="1"
maxOccurs="1"/>
<xs:element name="date" type="xs:string" minOccurs="0"
maxOccurs="1"/>
<xs:element name="description" type="xs:string" minOccurs="0"
maxOccurs="unbounded"/>
```

Identifikace

```
<xs:complexType name="identifierDefinition">
  <xs:simpleContent>
    <xs:extension base="stringPlusLanguage">
      <xs:attribute name="displayLabel" type="xs:string"/>
      <xs:attribute name="type">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:enumeration value="ORCID"/>
            <xs:enumeration value="researchID"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
      <xs:attribute name="typeURI" type="xs:anyURI"/>
      <xs:attribute name="invalid" fixed="yes"/>
      <xs:attribute name="altRepGroup" type="xs:string"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>

<xs:complexType name="stringPlusLanguage">
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attributeGroup ref="languageAttributeGroup"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>

<xs:attributeGroup name="languageAttributeGroup">
  <xs:attribute name="lang" type="xs:string"/>
  <xs:attribute ref="xml:lang"/>
  <xs:attribute name="script" type="xs:string"/>
  <xs:attribute name="transliteration" type="xs:string"/>
</xs:attributeGroup>
```

Schéma vazby na digitální objekt (reference citace)

```
<xs:element name="relatedItem" type="relatedItemDefinition"/>
<xs:complexType name="relatedItemDefinition">
  <xs:group ref="modsGroup" minOccurs="0" maxOccurs="unbounded"/>
  <xs:attribute name="type">
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="preceding"/>
        <xs:enumeration value="succeeding"/>
        <xs:enumeration value="original"/>
        <xs:enumeration value="host"/>
        <xs:enumeration value="constituent"/>
        <xs:enumeration value="series"/>
        <xs:enumeration value="otherVersion"/>
        <xs:enumeration value="otherFormat"/>
        <xs:enumeration value="isReferencedBy"/>
        <xs:enumeration value="references"/>
        <xs:enumeration value="reviewOf"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
  <xs:attribute name="otherType" type="xs:string"/>
  <xs:attribute name="otherTypeAuth" type="xs:string"/>
  <xs:attribute name="otherTypeAuthURI" type="xs:string"/>
  <xs:attribute name="otherTypeURI" type="xs:string"/>
  <xs:attribute name="displayLabel" type="xs:string"/>
  <xs:attribute name="ID" type="xs:ID"/>
</xs:complexType>
<xs:element name="identifier" type="identifierDefinition"/>
<!-- -->
<xs:complexType name="identifierDefinition">
  <xs:simpleContent>
    <xs:extension base="stringPlusLanguage">
      <xs:attribute name="displayLabel" type="xs:string"/>
      <xs:attribute name="type">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:enumeration value="ISBN"/>
            <xs:enumeration value="ISSN"/>
            <xs:enumeration value="DOI"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
      <xs:attribute name="typeURI" type="xs:anyURI"/>
      <xs:attribute name="invalid" fixed="yes"/>
      <xs:attribute name="altRepGroup" type="xs:string"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
```

H. XML záznam

```
<?xml version="1.0" encoding="UTF-8"?>
<msod>
  <!-- název díla -->
  <dc:title>
    Internet of Things (IoT) in Agriculture – Selected Aspects
  </dc:title>
  <!-- typ -->
  <dc:type>text</dc:type>
  <!-- datum vydání -->
  <dc:date>31.3.2016</dc:date>
  <!-- souhrn -->
  <dc:description>
    Article analyzes chosen aspects of Internet of Things (IoT) in
    general and in regards t.....
  </dc:description>
  <!-- identifikace -->
  <mods:identifier type="DOI">
    10.7160/aol.2016.080108
  </mods:identifier>
  <!-- hodnocení -->
  <rating>8</rating>
  <!-- klíčová slova AGROVOC -->
  <dc:subject xsi:type="dcterms:AGROVOC">internet</dc:subject>
  <dc:subject xsi:type="dcterms:AGROVOC">
    precision agriculture
  </dc:subject>
  <!-- další klíčová slova -->
  <!-- identifikace autorů -->
  <lom:lifeCycle>
    <lom:contribute >
      <lom:role>
        <lom:source>LOMv1.0</lom:source>
        <lom:value>Author</lom:value>
      </lom:role>
      <mods:identifier type="ORCID">
        orcid.org/0000-0001-7128-1071
      </mods:identifier>
    </lom:contribute >
  </lom:lifeCycle>
  <!-- další autoři -->
  <!-- identifikace reference -->
  <mods:relatedItem type="references">
    <mods:identifier type="doi">
      10.1109/COMST.2015.2444095
    </mods:identifier>
    <mods:part>
      <mods:extent unit="pages">
        <mods:start>2347</mods:start>
        <mods:end>2348</mods:end>
      </mods:extent>
    </mods:part>
  </mods:relatedItem>
  <!-- identifikace reference -->
```

```
<!-- idnetifikace citace -->  
<mods:relatedItem type="isReferedBy">  
  <mods:identifier type="doi">  
    10.1109/ICOS.2016.7881983  
  </mods:identifier>  
</mods:relatedItem>  
<!-- identifikace citace -->  
</msod>
```

I. Publikační činnost autora

Články J_{imp}

ŠIMEK, P. – VANĚK, J. – **STOČES, M.** – JAROLÍMEK, J. – PAVLÍK, J. Mobile accessibility expense analysis of the agrarian WWW portal. *Agricultural Economics*, 2017, roč. 2017, č. 5, s. 197–203. ISSN 1805-9295.

JAROLÍMEK, J. – VANĚK, J. – JEŽEK, M. – MASNER, J. – **STOČES, M.** The telemetric tracking of wild boar as a tool for field crops damage limitation. *Plant, Soil and Environment*, 2014, roč. 60, č. 9, s. 418-425. ISSN: 1214-1178.

ŠIMEK, P. – VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** – VOGELTANZOVÁ, T. Using metadata formats and AGROVOC thesaurus for data description in the agrarian sector. *Plant, Soil and Environment*, 2013, roč. 59, č. 8, s. 378-384. ISSN: 1214-1178.

Články J_{sc}

JAROLÍMEK, J. – **STOČES, M.** – MASNER, J. – VANĚK, J. – ŠIMEK, P. – PAVLÍK, J. – RAJTR, J. User-Technological Index of Precision Agriculture. *AGRIS on-line Papers in Economics and Informatics*, 2017, roč. 9, č. 1, s. 69-75. ISSN: 1804-1930.

STOČES, M. – VANĚK, J. – MASNER, J. – PAVLÍK, J. Internet of Things (IoT) in Agriculture - Selected Aspects. *AGRIS on-line Papers in Economics and Informatics*, 2016, roč. 8, č. 1, s. 83-88. ISSN: 1804-1930.

PAVLÍK, J. – VANĚK, J. – **STOČES, M.** Software tools for movement visualization in agrarian sector. *AGRIS on-line Papers in Economics and Informatics*, 2015, roč. 7, č. 2, s. 68-76. ISSN: 1804-1930.

STOČES, M. – MASNER, J. – JAROLÍMEK, J. Mitigation of Social Exclusion in Regions and Rural Areas – E-learning with Focus on Content Creation and Evaluation. *AGRIS on-line Papers in Economics and Informatics*, 2015, roč. 7, č. 4, s. 143-150. ISSN: 1804-1930.

MASNER, J. – VANĚK, J. – **STOČES, M.** Spatial Data Monitoring and Mobile Applications – Comparison of Methods for Parsing JSON in Android Operating System. *AGRIS on-line Papers in Economics and Informatics*, 2014, roč. VI, č. 1, s. 37-46. ISSN: 1804-1930.

ŠIMEK, P. – **STOČES, M.** – VANĚK, J. Mobile Access to Information in the Agrarian Sector. *AGRIS on-line Papers in Economics and Informatics*, 2014, roč. VI, č. 2, s. 89-96. ISSN: 1804-1930.

ŠIMEK, P. – **STOČES, M.** – VANĚK, J. – JAROLÍMEK, J. – MASNER, J. – HRBEK, I. Using of Automatic Metadata Providing. *AGRIS on-line Papers in Economics and Informatics*, 2013, roč. V, č. 4, s. 189-197. ISSN: 1804-1930.

JAROLÍMEK, J. – **STOČES, M.** Non-indigenous and Endangered Animal Record Keeping in the Czech Republic. *AGRIS on-line Papers in Economics and Informatics*, 2013, roč. V, č. 4, s. 115-120. ISSN: 1804-1930.

ŠIMEK, P. – VANĚK, J. – OČENÁŠEK, V. – **STOČES, M.** – VOGELTANZOVÁ, T. Using Metadata Description for Agriculture and Aquaculture Papers. *AGRIS on-line Papers in Economics and Informatics*, 2012, roč. IV, č. 4, s. 79-90. ISSN: 1804-1930.

VANĚK, J. – **STOČES, M.** – ŠIMEK, P. – HRBEK, I. Information Support of Regions and Possibilities of Its Further Development. *AGRIS on-line Papers in Economics and Informatics*, 2012, roč. IV, č. 3/2012, s. 71-78. ISSN: 1804-1930.

KÁNSKÁ, E. – ŠIMEK, P. – VOGELTANZOVÁ, T. – **STOČES, M.** Social Network as an Integration Tool in Rural Areas. *AGRIS on-line Papers in Economics and Informatics*, 2011, roč. 3, č. 1/2011, s. 53-60. ISSN: 1804-1930.

ŠIMEK, P. – VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** – VOGELTANZOVÁ, T. New Version of the AGRIS Web Portal – Overcoming the Digital Divide by Providing Rural Areas with Relevant Information. *AGRIS on-line Papers in Economics and Informatics*, 2011, roč. III, č. 4, s. 71-78. ISSN: 1804-1930.

Článek J_{rec}

ŠIMEK, P. – JAROLÍMEK, J. – VANĚK, J. – **STOČES, M.** – HRBEK, I. Implementace mechanismu protokolu OAI-PMH pro poskytování obsahu elektronického vědeckovýzkumného časopisu. *Systémová integrace*, 2011, roč. 18, č. 4, s. 55-65. ISSN: 1210-9479.

Články ve sborníku D

JAROLÍMEK, J. – **STOČES, M.** – VANĚK, J. IoT networks. In *AGRARIAN PERSPECTIVES XXV.: GLOBAL AND EUROPEAN CHALLENGES FOR FOOD PRODUCTION, AGRIBUSINESS AND THE RURAL ECONOMY 14.09.2016, Prague*. Prague: Czech University of Life Sciences Prague, Faculty of Economics and Management, 2016. s. 140-146.

- VANĚK, J. – OČENÁŠEK, V. – **STOČES, M.** – MASNER, J. Adoption of modern ICT in regional perspective – situation in the Czech Republic. In *GLOBALIZATION AND ITS SOCIO-ECONOMIC CONSEQUENCES 05.10.2016, Žilina, SK*. Žilina, SK: GEORG, Bajzova 11, 010 01 Zilina, Slovak Republic, 2016. s. 2291-2298.
- JAROLÍMEK, J. – ULMAN, M. – VANĚK, J. – MASNER, J. – **STOČES, M.** Use of ICT in education for mitigation of social exclusion. In *Future Communication, Information and Computer Science 22.05.2014, Beijing, China*. : CRC Press, 2015. s. 1-10.
- MASNER, J. – **STOČES, M.** – VANĚK, J. Spatial and activity data-collection and visualization in WWW environment . In *Future Communication, Information and Computer Science 22.05.2014, Beijing*. Beijing: CRC Press Taylor & Francis Group, 2015. s. 97-100.
- ULMAN, M. – ŠILEROVÁ, E. – MASNER, J. – **STOČES, M.** – ŠIMEK, P. – BENDA, P. Open data in regions from the users' perspective: an analytical study. In *Proceedings of the 7th International Conference on Information and Communication Technologies in Agriculture, Food and Environment 17.09.2015, Kavala, Greece*. Kavala, Greece: , 2015. s. 902-911.
- STOČES, M.** – BENDA, P. – ULMAN, M. Open data and rural development. In *Agrarian Perspectives XXIV. – Global Agribusiness and Rural Economy 16.09.2015, Prague*. Praha: Czech University of Life Sciences Prague , 2015. s. 399-406.
- STOČES, M.** – VANĚK, J. – MASNER, J. – JAROLÍMEK, J. Mobile application development options for news and information portals. In *Future Communication, Information and Computer Science 22.05.2014, Beijing*. Beijing: CRC press Taylor & Francis group, 2015. s. 111-114.
- ŠIMEK, P. – **STOČES, M.** – VANĚK, J. – MASNER, J. Mobile accessibility of information sources in the areas of agriculture, forestry, water management, food industry and rural development. In *Agrarian Perspectives XXIV. – Global Agribusiness and Rural Economy 16.09.2015, Prague*. Praha: Czech University of Life Sciences Prague, Faculty of Economics and Management, 2015. s. 440-446.
- STOČES, M.** – MASNER, J. – JAROLÍMEK, J. – ŠIMEK, P. – VANĚK, J. – ULMAN, M. Cross-platform user interface of e-learning applications. In *Proceedings of the 11th International Conference on Mobile Learning 2015, ML 2015 14.03.2015, Madeira*. Madeira; Portugal: IADIS, 2015. s. 135-138.
- MASNER, J. – **STOČES, M.** Web application for Collection and visualization of positional data - Wild boars and deer monitoring in Czech Republic. In *Agrarian Perspectives - Proceedings of the 22th Scientific Conference: Development trends in agribusiness 17.09.2013, Czech University of Life Sciences Prague, CR*. Prague, CR: CULS Prague, Faculty of Economics and Management, 2013. s. 141-147.
- VANĚK, J. – ŠIMEK, P. – **STOČES, M.** – JAROLÍMEK, J. – OČENÁŠEK, V. – BROŽOVÁ, I. Ecological farming - possibilities of presentation by means of map portal. In *6th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA 2013) 19.09.2013, Kerkyra, Greece*. : Procedia Technology, 2013. s. 561-565.
- STOČES, M.** – ŠIMEK, P. – JAROLÍMEK, J. – VANĚK, J. Metadata Providing from Content of Local Archives for Central Agriculture and Aquaculture Repository. In *AFITA/WCCA 2012 Asian Federation for Information Technology in Agriculture / World Congress on Computers in Agriculture 03.09.2012, Taipei, Taiwan*. Taipei, Taiwan: , 2012. s. 89-89.
- VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** – ŠIMEK, P. – KÁNSKÁ, E. Regional reports and presentations of regions in the Czech Republic . In *Proceedings of HAICTA 2011 08.09.2011, Skiathos; Greece*. : CEUR Workshop Proceedings, 2011. s. 667-674.
- ŠIMEK, P. – JAROLÍMEK, J. – VANĚK, J. – **STOČES, M.** Implementation of metadata harvesting of scientific and scholarly research journal's content. In *Proceedings of HAICTA 2011 08.09.2011, Skiathos; Greece*. : CEUR Workshop Proceedings, 2011. s. 827-834.
- Software**
- PAVLÍK, J. – JUNEK, P. – VANĚK, J. – JAROLÍMEK, J. – ŠIMEK, P. – MASNER, J. – **STOČES, M.** – GALLIE, M. Video Event Analyser for UX , Video Event Analyser for UX , UX; event; video; analyzer, 2016, JC - Počítačový hardware a software, VEAUX 1.0, <http://www.agris.cz/kit/veaux>.
- MASNER, J. – ŠIMEK, P. – VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** Electronic learning mobile application, Electronic learning mobile application, iOS; mobile application; e-learning; learning materials, 2016, JC - Počítačový hardware a software, ELMA 1.0, <http://www.agris.cz/kit/elma>.
- STOČES, M.** – JAROLÍMEK, J. – VANĚK, J. – ŠIMEK, P. – MASNER, J. Click Tracking Data Collector for UX, Click Tracking Data Collector for UX, web browser; tracking; user experience, 2016, JC - Počítačový hardware a software, CTDCUX 1.0, <http://www.agris.cz/kit/ctdcux>.

MASNER, J. – ŠIMEK, P. – **STOČES, M.** – VANĚK, J. – JAROLÍMEK, J. Nativní mobilní aplikace Agrárního WWW portálu AGRIS pro platformu Android, Agrarian WWW portal AGRIS native mobile application for Android platform, agris, android, aplikace, google play, 2015, JC - Počítačový hardware a software, Agris mobile v. 1.0, <http://www.agris.cz/kit/agm>.

STOČES, M. – ŠIMEK, P. – JAROLÍMEK, J. – VANĚK, J. PDF Cataloguer (PDFCat 1.0), PDF Cataloguer (PDFCat 1.0), PDF document, cataloging, DB server, indexing, key words, 2012, GA - Zemědělská ekonomie, PDFCat 1.0, www.agris.cz/kit/pdfcat.

ŠIMEK, P. – **STOČES, M.** – VANĚK, J. – JAROLÍMEK, J. – KARAS, L. Metadata Library (MetaLib 1.0) , Metadata Library (MetaLib 1.0) , Repository, OAI-PMH, metadata, information database, search service, 2012, GA - Zemědělská ekonomie, , www.agris.cz/kit/ml.

JAROLÍMEK, J. – VANĚK, J. – ŠIMEK, P. – **STOČES, M.** Hierarchical Term Assigner (HTA 1.0), Hierarchical Term Assigner (HTA 1.0), Code list, AGROVOC, term, key word, AJAX , 2012, GA - Zemědělská ekonomie, HTA 1.0, www.agris.cz/kit/hta

JAROLÍMEK, J. – MASNER, J. – **STOČES, M.** – ULMAN, M. – ŠIMEK, P. – VANĚK, J. Ground Station Harvester (GSH 1.0), Ground Station Harvester (GSH 1.0), GPS, positional data, ground station, database, SW, 2012, GA - Zemědělská ekonomie, GSH 1.0, www.agris.cz/kit/gsh.

MASNER, J. – JAROLÍMEK, J. – ULMAN, M. – ŠIMEK, P. – **STOČES, M.** – VANĚK, J. GPS Trace Route 1.0, GPS Trace Route 1.0, Positional data, GPS, polygon, quick hull, 2012, GA - Zemědělská ekonomie, GPS Trace Route 1.0, www.agris.cz/kit/tr.

VANĚK, J. – JAROLÍMEK, J. – ŠIMEK, P. – **STOČES, M.** DIT Questionnaire v 1.0, DIT Questionnaire v 1.0, Questionnaire, survey, software, respondent, 2012, GA - Zemědělská ekonomie, DIT Questionnaire v 1.0 , www.agris.cz/kit/ditq.

ŠIMEK, P. – VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** Agrarian Prices Harvester (APH 1.0), Agrarian Prices Harvester (APH 1.0), commodity exchange, pricing information, CMS, software component, 2012, GA - Zemědělská ekonomie, APH 1.0, www.agris.cz/kit/aph.

VANĚK, J. – ŠIMEK, P. – **STOČES, M.** – JAROLÍMEK, J. Agrarian Content Management System (ACMS 1.0), Agrarian Content Management System (ACMS 1.0), CMS, XML, themed cases, administration, aggregation, 2012, GA - Zemědělská ekonomie, ACMS 1.0, www.agris.cz/acms.

ŠIMEK, P. – JAROLÍMEK, J. – VANĚK, J. – BARTOŠKA, J. – **STOČES, M.** – KÁNSKÁ, E. ConfMaS 1.0 – Conference Management System, ConfMaS 1.0 – Conference Management System, conference, management system, web application, software, 2011, JC - Počítačový hardware a software, ConfMaS 1.0 , <http://ipc.pef.czu.cz/kit/ConfMaS>.

ŠIMEK, P. – VANĚK, J. – JAROLÍMEK, J. – **STOČES, M.** Agrární WWW portál Agris 5.0, Agrarian WWW portal Agris 5.0, portal, the World Wide Web, information, resource, rural development, rural, regions, agriculture, food, 2011, JC - Počítačový hardware a software, AGRIS 5.0, <http://www.agris.cz>, Informace k SW AGRIS jsou standardně k dispozici na: <http://ipc.pef.czu.cz/kit/agris>.

Ostatní výsledky, které nelze zařadit do žádného z výše uvedených druhů výsledku

JAROLÍMEK, J. – **STOČES, M.** – VANĚK, J. – ULMAN, M. Technology Support for Game Monitoring As a Tool for Damages Reduction of Field Crops. 2016, International Society of Precision Agriculture: International Conference on Precision Agriculture 13th (2016), International Scientific Conference, St. Louis, MI, USA.

ULMAN, M. – **STOČES, M.** – JAROLÍMEK, J. – ŠIMEK, P. Open data for food quality and food safety: a case study of the Czech Republic. 2016, International Society of Precision Agriculture: International Conference on Precision Agriculture 13th (2016), International Scientific Conference, St. Louis, MI, USA.

ULMAN, M. – JAROLÍMEK, J. – ŠIMEK, P. – MASNER, J. – VANĚK, J. – **STOČES, M.** EGOVERNMENT IN THE CZECH REPUBLIC – DO DATA RUN OR CITIZENS?. In *PROCEEDINGS OF THE 13th INTERNATIONAL CONFERENCE e-Society 2015 14.03.2015, Funchal, Madeira*. : IADIS Press, 2015. s. 255-259.

VANĚK, J. – ŠIMEK, P. – **STOČES, M.** – ULMAN, M. – MASNER, J. – JAROLÍMEK, J. DIGITAL DIVIDE AND ICT DEVELOPMENT IN RURAL AREAS IN THE CZECH REPUBLIC. In *PROCEEDINGS OF THE 13th INTERNATIONAL CONFERENCE e-Society 2015 14.03.2015, Funchal, Madeira - Portugal*. Portugal: IADIS Press, 2015. s. 331-334.

MASNER, J. – **STOČES, M.** Sběr a zobrazení pozičních dat pomocí webové aplikace. In *Think Together 2014, Doktorská vědecká konference 03.02.2014, Praha*. Praha: Provozně ekonomická fakulta, Česká zemědělská univerzita v Praze, 2014.

JAROLÍMEK, J. – ULMAN, M. – **STOČES, M.** Sběr a interpretace pozičních dat. In *Sika japonský - škody, perspektiva, metody a technologie výzkumu, možnosti zpracování dat 01.06.2012, Dubina, Karlovy Vary*. Dubina, Karlovy Vary: Česká lesnická společnost. o.s., 2012. s. 46-46.

JAROLÍMEK, J. – ŠIMEK, P. – VANĚK, J. – **STOČES, M.** National Registers – eGovernment pillar in the Czech Republic. In *3rd World Conference on Information Technology 2012 14.11.2012, Barcelona*. Barcelona: Near East University, 2012. s. 99-100.

ŠIMEK, P. – **STOČES, M.** – JAROLÍMEK, J Usability of Open Source Content Management Systems in Agrarian Sector. In *XVI. Informační systémy v zemědělství a lesnictví*, 2010, Praha, ČZU v Praze, ISBN 978-80-213-2087-1