



národní
úložiště
šedé
literatury

Metodika zpracování, dlouhodobého uchování a zpřístupnění šedé literatury v ČR na příkladu Národního úložiště šedé literatury

Petra Pejšová a kol.

Realizováno z programového projektu DC - Zpřístupnění a ochrana kulturních, uměleckých a vědeckých zdrojů (2008 – 2011) financovaného MK ČR v rámci projektu „Digitální knihovna pro šedou literaturu – funkční model a pilotní realizace“, identifikační kód: DC08P02OUK007

Praha
Národní technická knihovna
2011

1 CÍL METODIKY.....	3
2 ÚVOD.....	4
2.1 Teoretický rámec pro budování systémů zpřístupňujících a dlouhodobě archivujících digitální dokumenty.....	4
2.2 Šedá literatura.....	4
2.3 Cílová skupina metodiky.....	5
2.4 Struktura metodiky.....	5
3 DEFINOVÁNÍ REPOZITÁŘE	6
3.1 Legislativní rámec.....	6
3.2 Určení obsahu digitálního repozitáře.....	12
3.3 Výběr software pro digitální repozitář.....	17
4 ZPRACOVÁNÍ ŠEDÉ LITERATURY.....	19
4.1 Metadatový formát.....	19
4.2 Šablony pro šedou literaturu.....	27
4.3 Vkládání dat do digitálního repozitáře.....	30
5 DLOUHODOBÉ UCHOVÁNÍ A ZPŘÍSTUPNĚNÍ ŠEDÉ LITERATURY.....	33
5.1 Doporučené formáty a dlouhodobá udržitelnost digitálních dokumentů.....	33
5.2 Persistentní identifikátor	36
5.3 Zpřístupňování a předávání dat	42
5.4 Předávání dat do národních a mezinárodních center.....	49
6 SROVNÁNÍ NOVOSTI POSTUPŮ.....	53
7 POPIS UPLATNĚNÍ CERTIFIKOVANÉ METODIKY.....	53
8 POUŽITÁ LITERATURA.....	54
9 SEZNAM PUBLIKACÍ	57
10 SEZNAM ZKRATEK.....	60

1 Cíl metodiky

Metodika předkládá postup zpracování, dlouhodobého uchování a zpřístupňování šedé literatury v České republice. Cílem metodiky je poskytnout návod pro producenty šedé literatury, jak budovat digitální repozitáře a přitom dodržet standardy pro dlouhodobé zpřístupnění jejich digitálních dokumentů a pro spolupráci s agregátory na národní i mezinárodní úrovni. Veškeré postupy a doporučení byly ověřeny v projektu Digitální knihovna pro šedou literaturu – funkční model a pilotní realizace, který řešila Národní technická knihovna společně s Vysokou školou ekonomickou v Praze v letech 2008-2011 díky finanční podpoře z programu Zpřístupnění a ochrana kulturních, uměleckých a vědeckých zdrojů Ministerstvo kultury České republiky.

2 Úvod

2.1 Teoretický rámec pro budování systémů zpřístupňujících a dlouhodobě archivujících digitální dokumenty

Cílem metodiky není vytvořit obecný návod na budování digitálního repozitáře, ale ukázat praktické postupy při vytváření digitálního uložení pro šedou literaturu a vymezit jeho nastavení způsobem, který odpovídá podmínkám pro připojení do národních a mezinárodních systémů šedé literatury.

Při budování digitálních repozitářů je proto potřeba vycházet nejen z této metodiky, ale také z obecných mezinárodně uznávaných standardů a norem pro tvorbu digitálních repozitářů.

Doporučené dokumenty:

Pro metadata a The Dublin Core metadata element set

ISO 15836:2003 [online]. © ISO 2011. [cit. 2011-11-18]. Dostupné z [www: <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37629>](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37629)

Pro dlouhodobou ochranu a zpřístupnění

Consultative Committee for Space Data Systems. 2002. *Reference model for an open archival information system (OAIS) : CCSDS 650.0-B-1* [online]. Washington (DC, USA) : CCSDS, January 2002 [cit. 2011-11-10]. Blue Book, Issue 1. Dostupné z WWW: <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>

Pro dlouhodobou ochranu

PREMIS [online]. The Library of Congress, November 17, 2011. [cit. 2011-11-18]. Dostupné z [www: <http://www.loc.gov/standards/premis/>](http://www.loc.gov/standards/premis/)

Pro zajištění interoperability

The Open Archives Initiative Protocol for Metadata Harvesting [online]. [cit. 2011-11-18]. Dostupné z [www: <http://www.openarchives.org/OAI/openarchivesprotocol.html>](http://www.openarchives.org/OAI/openarchivesprotocol.html)

Pro zajištění auditu

Trustworthy Repository Audit & Certification: Kriteria and Checklist [online]. Version 1.0, OCLC, February 2007. [cit. 2011-11-18]. Dostupné z [www: <http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf>](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)

Digital Repository Audit Method Based on Risk Assessment [online]. © Digital Curation Centre and Digital Preservation Europe. 1 February 2008. [cit. 2011-11-18]. Dostupné z [www: <http://www.repositoryaudit.eu/>](http://www.repositoryaudit.eu/)

2.2 Šedá literatura

Existuje několik definic šedé literatury, z nichž celosvětově nejpoužívanější je definice vznikla na konferenci 3rd International Conference on Grey Literature (dále jen GL) v Luxemburgu v roce 1997 a byla následně rozšířena na konferenci GL6

(6th International Conference on Grey Literature) v New Yorku v roce 2004¹. Volně přeloženo: *Šedá literatura jsou informace produkované na všech úrovních vládních, akademických, obchodních a průmyslových institucí jak v elektronické tak v tištěné podobě, které neprošly standardním vydavatelským procesem či nejsou distribuovány do standardní prodejní sítě, tj. vydávané institucemi, jejichž hlavní činností není vydavatelská činnost.*

Pro vědecké účely je nezbytné používat vedle tradičně publikovaných zdrojů též šedou literaturu, protože je v ní publikováno mnoho cenných údajů. V tomto případě se jedná především o zprávy ze všech úrovní vládních a akademických orgánů, výroční zprávy podniků, závěrečné zprávy z projektů, konferenční materiály, preprinty apod. Diskuse se otevírá nad informacemi typu článků z blogů, telefonních rozhovorů, diskusí, e-mailů apod., které jistě pomůžou dotvořit komplexní obraz o daném tématu, ale je nutné je před použitím pečlivě ověřit. Diskutabilní je i sběr a dlouhodobé uchování takovéto neověřené literatury. Při sběru šedé literatury by se proto měl klást důraz na „důvěryhodnost“ informací, kterou zajišťují buď samotní producenti jako jsou vědecké instituce, vysoké školy, státní správa atd., anebo výběrem těch druhů šedé literatury, které již prošly ověřovacím procesem, jako jsou závěrečné zprávy, konferenční materiály apod. (PEJŠOVÁ, 2010)

2.3 Cílová skupina metodiky

Vzrůstající potřeba přístupu ke sbírkám šedé literatury z oblasti vědy, výzkumu a vzdělávání vede ke koordinovanému sběru a budování sbírek na všech úrovních. Tato metodika je určena pro samotné producenty šedé literatury. Cílem metodiky je poskytnout této cílové skupině návod a legislativní rámec, které jim pomohou při budování lokálních digitálních repozitářů tak, aby zajistili dlouhodobé uchování a zpřístupnění dat a aby jejich repozitáře splňovaly standardy potřebné pro národní i mezinárodní integraci.

Do této cílové skupiny patří především tyto organizace:

- Výzkumné ústavy v ČR a pracoviště AV ČR, v.v.i.
- Grantové agentury
- Vysoké školy
- Knihovny, muzea a další paměťové instituce
- Orgány státní správy, samosprávy a další subjekty veřejné správy
- Podniky, firmy, obchodní organizace a společnosti

2.4 Struktura metodiky

Metodika je rozdělena do tří hlavních oblastí, kterými jsou 1) definování repozitáře, 2) zpracování dat a 3) dlouhodobé uchování a zpřístupnění digitálních dokumentů. Tyto tři hlavní oblasti jsou dále děleny na části. Každá část obsahuje úvod, řešení hlediska, popis, příklad, doporučení a, pokud existují, normy, standardy a legislativu.

¹ "Information produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishing i.e. where publishing is not the primary activity of the producing body." (Luxembourg, 1997 - Expanded in New York, 2004, dostupné na WWW <<http://www.greynet.org/index.html>>)

3 Definování repozitáře

3.1 Legislativní rámec

Důležitým aspektem je dodržování českých právních norem. Legislativní rámec pro zpracování, dlouhodobé uchování a zpřístupňování digitálních zdrojů šedé literatury je tvořen několika různými právními normami. V české legislativě nenajdeme žádný specifický právní předpis, který by se věnoval pouze šedé literatuře.

Komplexně se na všechny druhy šedé literatury vztahuje Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon). Šedá literatura je autorským dílem a provozovatel databáze² může plný text šedé literatury zveřejnit pouze se souhlasem vykonavatele majetkových autorských práv. Autorský zákon nabízí řešení v podobě licenčních smluv, které uzavírá provozovatel databáze se subjektem disponujícím majetkovými autorskými právy k dílu. Zde je třeba postupovat v souladu s obecným smluvním právem. Provozovatel databáze by neměl opomenout Zákona č. 101/2000 Sb., o ochraně osobních údajů (dále jen ZOOÚ) a ověření, zda je nutné registrovat databázi u Úřadu pro ochranu osobních údajů (dále jen ÚOOÚ).

Nejpřesněji definuje zacházení s konkrétním druhem šedé literatury Zákon č. 111/1998 Sb., o vysokých školách (dále jen ZVŠ), který specifikuje vysokoškolské kvalifikační práce (dále jen VŠKP). Na druhé straně téměř chybí legislativní úprava v oblasti výzkumu a vývoje v České republice. Zákon č. 130/2002 Sb., o podpoře výzkumu a vývoje (dále jen Zákon o podpoře VaV) se vyjadřuje k poskytování informací o výzkumu a vývoji v Informačním systému výzkumu a vývoje (dále jen IS VaV). Zpracování a zpřístupňování výročních zpráv řeší kromě zákona o účetnictví i mnohé další specifické právní normy.

Producent šedé literatury je při sběru šedé literatury limitován externími právními předpisy, ale k zajištění systematického a efektivního shromažďování šedé literatury je třeba vytvořit interní směrnici. Tato směrnice by měla upravovat proces odevzdávání, evidování, ukládání a zpřístupňování zaměstnaneckých nebo školních děl.

Které právní předpisy zohlednit při budování digitálních repozitářů:

- 1) Autorský zákon a jeho majetkovou složku,
- 2) Ochranu osobních údajů a registraci u Úřadu pro ochranu osobních údajů,
- 3) Zákon o vysokých školách a povinnost zveřejňovat vysokoškolské kvalifikační práce,
- 4) Zákon o podpoře výzkumu a vývoje upravující zveřejňování výsledků,
- 5) Zákon ukládající povinnost zveřejňovat výroční zprávy,
- 6) Zákon o svobodném přístupu k informacím,
- 7) vytvoření vlastních směrnic.

² V celé části Legislativní rámec bude místo „digitálního repozitáře“ používán termín „databáze“, který je převzat z Autorského zákona a používají ho i další zákony např. Zákon o vysokých školách

3.1.1 Autorský zákon

Šedá literatura je z právního hlediska relativně složitý pojem, kterému se specificky nevěnuje žádná platná česká právní norma. Legislativní úpravy týkající se šedé literatury je proto třeba hledat v několika různých právních normách. Až na malé výjimky není v zákonech upraven formální vzhled šedé literatury, ani způsoby archivace a zveřejnění.

Pro šedou literaturu jako celek platí Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon). Jakýkoliv druh šedé literatury je podle tohoto zákona považován za autorské dílo (§2 odst. 1, autorského zákona). K takovému dílu se váží autorská práva, která zahrnují osobnostní a majetkovou složku (§10 autorského zákona). Osobnostní složka autorských práv se váže bezprostředně k osobě autora, ale pro pořizování a vedení databáze je mnohem důležitější majetková složka autorských práv, která zahrnuje práva příslušné dílo užívat.

K tomu, aby mohl provozovatel databáze určitý dokument šedé literatury tímto způsobem užít, je třeba, aby k tomu disponoval příslušným oprávněním. To může být uděleno autorem či jiným vykonavatelem majetkových práv autorských (např. u zaměstnaneckých děl výzkumným ústavem, univerzitou, firmou, ministerstvem, úřadem, atd.). Zde je třeba upozornit, že autorská díla, která vznikla do 30. 11. 2000, se řídí podle Zákona č. 35/1965 Sb. o autorských právech, který upravuje problematiku zaměstnaneckých děl v § 17 takto: „Zaměstnavatel, do jehož předmětu činnosti náleží vydávat nebo jinak uveřejňovat díla, může vydat nebo jinak uveřejnit dílo svého zaměstnance vytvořené ke splnění povinností vyplývajících z pracovního poměru jen se svolením autora.“ Od 1. 12. 2000, kdy začal platit současný autorský zákon, jsou již majetková autorská práva k zaměstnaneckým dílům vyhrazena pouze zaměstnavateli.

Nevztahuje-li se na autorské dílo zákonná licence nebo volné užití, lze dílo užít jen na základě licenční smlouvy, kterou se subjektem vykonávajícím k dílu majetková autorská práva uzavře provozovatel databáze.³ Licence tak musí mít povahu dvoustranného právního úkonu a vztahuje se jen na digitální dokumenty, tj. na plné texty. Popisná metadata jsou veřejně dostupná bez nutnosti poskytnutí licence.

Licenční smlouva, která poskytne provozovateli právo dílo užít dohodnutým způsobem, je definována v autorském zákoně v odd. 1 § 46 – 55. Při sestavování licenční smlouvy je třeba respektovat ustanovení Zákona č. 513/1991 Sb., obchodního zákoníku. Pokud provozovatel databáze poskytuje několik druhů služeb, může si vytvořit několik vzorů licenčních smluv, jenž mu usnadní další komunikaci s producenty dat.⁴

Autorské dílo je tedy chráněno autorským zákonem nebo je možné využít některou z licencí Creative Commons (dále jen CC), což je soubor veřejných licencí, které přinášejí nové možnosti v oblasti publikování autorských děl, tj. posilují pozici autora při rozhodování, za jakých podmínek bude dílo veřejně zpřístupněno.⁵

3.1.2 Ochrana osobních údajů

Pořizovatel databáze je povinen si ověřit u Úřadu pro ochranu osobních údajů (ÚOOÚ), zda data shromažďovaná v databázi jsou nebo nejsou zpracováním osobních údajů. Pokud databáze zpracovává osobní případně citlivé údaje

³ Za takovou dohodu nelze považovat skutečnost, že autor vystavil své dílo na volně dostupném webu. (PEJŠOVÁ, 2010b, s. 68.)

⁴ Pro NUŠL jsou vzory licenčních smluv součástí POLČÁK, 2009.

⁵ <http://www.creativecommons.cz/>

(vysvětleno níže), musí pořizovatel databázi registrovat u ÚOOÚ a řídit se podle ustanovení Zákona č. 101/2000 Sb., o ochraně osobních údajů a o změně některých zákonů (dále jen ZOOÚ).⁶

Například pro metadata, která shromažďuje a zveřejňuje NUŠL, platí §18 odst.1, písm. b) ZOOÚ, kdy NUŠL shromažďuje osobní údaje podle zvláštního zákona, tj. Zákon č. 257/2001 Sb. o knihovnách a podmínkách provozování veřejných knihovnických a informačních služeb (knihovní zákon) a Zákon č. 480/2004 Sb., o některých službách informační společnosti a o změně některých zákonů (zákon o některých službách informační společnosti).

Osobním údajem je „jakákoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt údajů se považuje za určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu“ (§ 4 písm. a ZOOÚ)

Citlivý údaj je „osobní údaj vypovídající o národnostním, rasovém nebo etnickém původu, politických postojích, členství v odborových organizacích, náboženství a filozofickém přesvědčení, odsouzení za trestný čin, zdravotním stavu a sexuálním životě subjektu údajů a jakýkoliv biometrický nebo genetický údaj subjektu údajů.“ (§4 písm. b, ZOOÚ)

Pojmy osobní údaj a citlivý údaj na sebe navazují, neboť podle ZOOÚ je "...citlivým údajem osobní údaj..." (§ 4 písm. b). Lze proto říci, že citlivé údaje jsou podmnožinou osobních údajů. Porovnáním obou úprav lze říci, že úprava citlivých údajů dopadá obsahově na stejný okruh otázek, jako úprava týkající se osobních údajů. Lze proto soudit, že jde o poměr úpravy speciální (citlivé údaje) vůči úpravě obecné (osobní údaje) (TONER, 2003).

Osobní údaj je z obecného hlediska širší kategorie než citlivý osobní údaj. Tzn. jakýkoli citlivý osobní údaj je zároveň osobním údajem.

Osobní údaj je každý údaj, který vede k identifikaci konkrétní osoby. Z toho pak vyplývá, že například údaj "křestní jméno: Jan" rozhodně není osobním údajem, avšak ve spojení s dalšími údaji "příjmení: Novák a adresa: Veveří 24, Brno" se již o osobní údaj jedná. Je spousta problematických údajů, o kterých se vedou teoretické spory: emailová adresa, IP adresa atd. Takovými bezpečnými příklady, které se dají uvést, je jméno a příjmení a také adresa. Citlivý osobní údaj musí splňovat znaky osobního údaje a ještě navíc musí něco vypovídat o "o národnostním, rasovém nebo etnickém původu, politických postojích, členství v odborových organizacích, náboženství a filozofickém přesvědčení, odsouzení za trestný čin, zdravotním stavu a sexuálním životě subjektu údajů a genetický údaj subjektu údajů" (také biometrická informace). Například tedy údaj "zdravotní stav: HIV pozitivní" nebo "vyznání: katolík" není citlivým osobním údajem (neidentifikuje osobu). Jakmile ovšem kterýkoli z těchto údajů bude připojen k výše uvedeným "jméno: Jan příjmení: Novák adresa: Veveří 24, Brno" bude se jednat o citlivý osobní údaj.

3.1.3 Zákon o vysokých školách

Nejkonkrétněji definuje nakládání s šedou literaturou Zákon č. 111/1998 Sb., zákon o vysokých školách a o změně a doplnění některých dalších zákonů a Zákon č. 552/2005 Sb., kterým se mění ZVŠ, a některé další zákony. Tyto zákony upravují zveřejňování závěrečných VŠKP, tj. bakalářských, diplomových, rigorózních a disertačních prací (§47b ZVŠ). Jsou to díla, která vznikají v průběhu studia a jejichž

⁶ Povinnosti správce osobních údajů § 5 - § 11 ZOOÚ.

účelem je prokázat splnění studijních povinností nebo dosažení dostatečné úrovně odbornosti k udělení akademického titulu (PEJŠOVÁ, 2010b, s. 71). Mezi VŠKP řadíme i habilitační práce, které definuje ZVŠ takto:

„(3) Habilitační prací se rozumí:

- A písemná práce, která přináší nové vědecké poznatky, nebo
- B soubor uveřejněných vědeckých prací nebo inženýrských prací doplněný komentářem, nebo
- C tiskem vydaná monografie, která přináší nové vědecké poznatky, nebo
- D umělecké dílo nebo umělecký výkon nebo jejich soubor, kterým je například vynikající veřejná umělecká činnost.“ (§ 72 ZVŠ)

VŠKP představují významnou část šedé literatury. Do roku 2006 byly VŠKP dostupné v univerzitních knihovnách prezenčně v tištěné podobě pouze studentům a zaměstnancům univerzity na základě školní licence dle autorského zákona.⁷ Zákon 216/2006 Sb., kterým se mění autorský zákon, mimo jiné vymezil nové situace, kdy užitím díla nedochází k zásahu do autorského práva. VŠKP se týká § 37 Knihovní licence, ale ta ještě neopravňovala ke zpřístupnění práce volně prostřednictvím internetu.

Významnou legislativní změnu představoval Zákon č. 552/2005 Sb. vkládající do ZVŠ § 47 b). Ten univerzitám ukládá povinnost nevýdělečně zveřejňovat disertační, diplomové, bakalářské a rigorózní práce, u kterých proběhla obhajoba, včetně posudků oponentů a výsledku obhajoby. Tyto práce musí být též nejméně pět dní před obhajobou dostupné veřejnosti, je možné si z nich pořizovat výpisy, opisy nebo rozmnoženiny. Způsob, jakým jsou práce zveřejněny v databázi kvalifikačních prací, má stanovit vnitřní předpis vysoké školy. Vnitřní předpisy vysokých škol se liší (viz příklady směrnic o VŠKP některých vysokých škol níže). Autor odevzdáním své práce souhlasí s takovýmto zveřejněním bez ohledu na výsledek obhajoby. ZVŠ neupravuje zveřejňování prací habilitačních (PEJŠOVÁ, 2010b, s. 55 – 65).

3.1.4 Zákon o podpoře výzkumu a vývoje

Šedá literatura vznikající v oblasti výzkumu a vývoje v České republice téměř postrádá legislativní normy, podle nichž by bylo možné jakkoliv postupovat při sběru, archivaci a zpřístupňování digitálních zdrojů šedé literatury. Zákon č. 130/2002 Sb., o podpoře výzkumu a vývoje z veřejných prostředků a o změně některých souvisejících zákonů (zákon o podpoře VaV) se vyjadřuje k poskytování informací o výzkumu a vývoji v § 30 až v § 32. Tato ustanovení se ovšem zabývají zveřejňováním povinností v informačním systému výzkumu a vývoje. V § 32 je popsán obsah IS VaV, který zahrnuje Centrální evidenci projektů výzkumu a vývoje, Centrální evidenci výzkumných záměrů a Rejstřík informací o výsledcích (dále jen RIV), které nezahrnují žádné digitální dokumenty (plné texty). O zveřejňování digitálních dokumentů např. závěrečných zpráv z projektů, výzkumných zpráv, aj. není v zákoně o podpoře VaV žádná zmínka, přestože jsou to výstupy projektů hrazených z veřejných prostředků.

3.1.5 Zákony ukládající povinnost zveřejňovat výroční zprávy

Výroční zpráva, jejímž účelem je podle Zákona č. 563/1991 Sb., o účetnictví (zákon o účetnictví) uceleně, vyváženě a komplexně informovat o vývoji společností, jejich výkonnosti, činnosti a stávajícím hospodářském postavení (§ 21 zákona o účetnictví). V § 21a jsou konkrétně upraveny způsoby zveřejnění výroční zprávy. Státní správa

⁷ Podepisovala-li škola se studenty licenční smlouvu, bylo možné půjčovat tyto práce i širší veřejnosti. (PEJŠOVÁ, 2010b, s. 56)

předkládá výroční zprávy dle Zákona č. 106/1999 Sb., o svobodném přístupu k informacím (viz níže). Povinnost předkládat výroční zprávu mají příspěvkové organizace ministerstev, měst či krajů, neboť hospodaří s financemi od svého zřizovatele. Dále vypracovávají výroční zprávu obecně prospěšné společnosti zřízené podle Zákona č. 248/1995 Sb., o obecně prospěšných společnostech, v platném znění. Podobně také tvoří výroční zprávu nadace a nadační fondy, které pracují podle Zákona číslo 227/1997 Sb., o nadacích a nadačních fondech, v platném znění. Společníci, představenstvo, respektive akcionáři akciové společnosti, také mají podle obchodního zákoníku právo seznámit se s výroční zprávou akciové společnosti. Výroční zprávu musejí předkládat ekonomické subjekty, jejichž účetnictví ověřuje auditor.

3.1.6 Zákon o svobodném přístupu k informacím

K problematice šedé literatury se vztahuje také Zákon č. 106/1999 Sb., o svobodném přístupu k informacím, který upravuje podmínky práva svobodného přístupu k informacím a stanoví základní podmínky, za nichž jsou informace poskytovány (§ 1). Povinnost poskytovat informace vztahující se k jejich působnosti mají podle tohoto zákona státní orgány a orgány územní samosprávy. Zveřejňovanou informací je taková informace, která může být vždy znovu vyhledána a získána (§ 3, odst. 3). Povinné subjekty, tj. orgány státní správy a samosprávy, jsou povinny pro poskytování informací využít také dálkového přístupu (např. Internet) a může informace zveřejnit i dalšími způsoby, např. v databázi šedé literatury (§ 5 odst. 4). Již výše zmíněná výroční zpráva má obsahovat zprávu o činnosti v oblasti poskytování informací (§ 18).

3.1.7 Vlastní směrnice

Šedá literatura se v institucích obvykle shromažďovala v knihovně, kam se dostávala v tištěné podobě. Mnohé instituce měli rovněž své nakladatelství, které vydávalo sborníky, brožurky, studijní materiály, výroční publikace, apod. S rozvojem informačních technologií postupně ubývá tištěných publikací, které se řadí do oblasti šedé literatury, a mnohé instituce s finančních důvodů již vlastní nakladatelství neprovozují. Většina dokumentů zůstává v digitální formě. Instituce ať už z veřejného nebo komerčního sektoru jsou v současné době postaveny před řešením problému, jak tento zdroj informací shromažďovat, uchovávat a zpřístupnit i pro další používání. Řešení nabízí institucionální repozitář, který si buď instituce vede sama, nebo jej pro ni vede třetí osoba. Ke zdárnému naplňování institucionálního repozitáře je nezbytné vytvořit v rámci instituce vnitřní předpis, který bude upravovat proces odevzdávání dokumentů, evidování, ukládání a zpřístupňování zaměstnaneckých nebo školních děl.

Pod pojmem zaměstnanecké dílo rozumíme dílo, které „autor vytvořil ke splnění svých povinností vyplývajících z pracovněprávního či služebního vztahu k zaměstnavateli nebo z pracovního vztahu mezi družstvem a jeho členem“ (§ 58 autorského zákona). Také u zaměstnaneckého díla náleží autorovi osobnostní autorská práva, ale „není-li sjednáno jinak, zaměstnavatel vykonává svým jménem a na svůj účet autorova majetková práva“ (§ 58 autorského zákona). Například výroční zpráva výzkumného ústavu, která byla sepsána společně konkrétními vědeckými pracovníky, tak může být zveřejněna s autorskou doložkou označující za autora pouze výzkumný ústav (PEJŠOVÁ, 2010b, s. 73 – 75).

Školní díla (§ 60 autorského zákona) vznikají v průběhu studia a jejich účelem je prokázat splnění studijních povinností nebo dosažení dostatečné úrovně odbornosti k udělení akademického titulu. Zde se jedná o nejrůznější formy průběžných studijních plnění, jako jsou např. seminární nebo semestrální práce, ale též VŠKP, na které se vztahuje současně autorský zákon a zákon o vysokých školách.

Směrnice by měla být v souladu s již platnými předpisy, případně upravit jejich platnost. Ve směrnici je nutné vymezit pojmy, které se budou v dokumentu vyskytovat. Například: autor, zaměstnavatel, zaměstnanec, zaměstnanecké nebo školní dílo, datum zveřejnění, atd. Neméně důležité je stanovení formálních náležitostí, které má dokument obsahovat. Toto je velmi důležité ve směrnících univerzit, které upravují odevzdávání VŠKP, z důvodů zachování jednotnosti. V interním dokumentu nesmí chybět formát, v němž má být práce uložena (např. Adobe PDF, TIF) a časový horizont pro odevzdání práce (např. do 1 měsíce od vytvoření dokumentu). Dále popis procesu odevzdání, adresa repozitáře a instrukce pro vložení do repozitáře. Případně je nutné zde uvést pověřené pracoviště, kam práce odevzdávat.

Jako příklad může sloužit *Směrnice 1/2011 ředitele NTK o publikační činnosti zaměstnanců NTK a právech k výsledkům VaVal* [online]. Národní technická knihovna. [cit. 2011-09-23]. Dostupná z [www: http://repozitar.techlib.cz/record/85](http://repozitar.techlib.cz/record/85) . Tato směrnice upravuje podmínky publikování zaměstnanců NTK, evidování, archivaci a zpřístupnění zaměstnaneckých děl a přístup k nim. Dalším příkladem mohou být směrnice některých vysokých škol.

Příklady směrnic o VŠKP některých českých vysokých škol

Směrnice rektora č. 120/2008 o vysokoškolských kvalifikačních pracích [online]. Ostravská univerzita v Ostravě. [cit. 2011-07-04]. Dostupný z [www: http://knihovna.osu.cz/dokumenty/smer_rek-2008120.pdf](http://knihovna.osu.cz/dokumenty/smer_rek-2008120.pdf)

Směrnice děkana č. 2/2009 (úprava odevzdání a zveřejňování vysokoškolských kvalifikačních prací) [online]. Vysoké učení technické v Brně. [cit. 2011-07-04]. Dostupný z [www: http://www.diplomky-online.cz/wp-images/smernice/2009-2.pdf](http://www.diplomky-online.cz/wp-images/smernice/2009-2.pdf)

Výnos rektora č. 4/2009. Metodika jednotné úpravy, zpracovávání, ukládání a zpřístupňování vysokoškolských kvalifikačních prací na AMU [online]. Akademie múzických umění v Praze. c2007-2011. [cit. 2011-07-04]. Dostupný z [www: http://www.amu.cz/amu/predpisy/vynosy-rektora/platne/vynos-rektora-c-4-2009](http://www.amu.cz/amu/predpisy/vynosy-rektora/platne/vynos-rektora-c-4-2009)

Jednotná formální úprava vysokoškolských kvalifikačních prací, jejich uložení a zpřístupnění [online]. Směrnice 5/2006. Univerzita Tomáše Bati ve Zlíně. [cit. 2011-07-04]. Dostupný z [www: http://web.utb.cz/cs/docs/smernice_r5_2006.pdf](http://web.utb.cz/cs/docs/smernice_r5_2006.pdf)

Doporučení:

Instituce jako provozovatel databáze musí mít právo užít dílo dle autorského zákona. To lze řešit buď licenční smlouvou, nebo tzv. veřejnými licencemi, které v České republice zastupují především Creative Commons. Provozovatel databáze je vždy povinen si ověřit u ÚOOÚ, zda data shromažďovaná v databázi jsou nebo nejsou zpracováním osobních údajů. Přestože zákon o podpoře výzkumu a vývoje neukládá institucím povinnost ukládat a zveřejňovat dokumenty vznikající v průběhu a na závěr výzkumu, je pro další úspěšný vývoj české vědy důležité, aby tuto šedou literaturu zveřejňovaly open access ve svých digitálních repozitářích. Zveřejnění prostřednictvím digitálního repozitáře se týká také výročních zpráv institucí, které v současné době jsou obvykle zveřejňovány na webových stránkách institucí, které

se často mění a nezaručují dlouhodobou dostupnost výročních zpráv. K výše zmíněným doporučením je nezbytné, aby si instituce vytvořily vlastní směrnice.

3.2 Určení obsahu digitálního repozitáře

Účelem této části metodiky je popsat způsob definování obsahu digitálního repozitáře. Definování obsahu při budování digitálních repozitářů má podobnou funkci jako akviziční politika jakékoliv klasické knihovny. Navíc je zde potřeba přihlídnout k technologickým aspektům budování těchto digitálních úložišť a právním aspektům, které jsou popsány v legislativním rámci. Jasně definování obsahu repozitáře je důležité jak pro správce a management, tak směrem k tvůrcům dat, ale i uživatelům.

Při budování digitálního repozitáře šedé literatury je poměrně náročné vymezit jeho obsah. Při jeho definování je potřeba mít na zřeteli více ukazatelů, na jejichž základě je pak vytvořena typologie dokumentů. Typologie dokumentů jasně určuje, které dokumenty budou zařazovány do úložiště. Dostupné typologie v oblasti šedé literatury se zaměřují pouze na druhy dokumentů. Při budování digitálního repozitáře je přitom potřeba k definování obsahu přistupovat z různých hledisek.

Při definování obsahu digitálních repozitářů je potřeba zohlednit následující hlediska:

- 1) druh dokumentu,
- 2) obsahové,
- 3) časové,
- 4) formu,
- 5) geografické,
- 6) stupeň zveřejnění,
- 7) jazykové.

3.2.1 Hledisko dle druhu dokumentu

Vymezení obsahu digitálního repozitáře z hlediska typologie dokumentu je částečně dáno již definicí samotného pojmu šedá literatura. Na konkrétní instituci pak záleží, které druhy šedé literatury produkuje, a následná potřeba zařadit je do digitálního repozitáře.

Charakteristika hlavních skupin druhů dokumentů šedé literatury

- **Vysokoškolské kvalifikační práce** - dokumenty, které jsou výsledkem studia jednotlivých studijních programů a oborů vysokých škola a vedou k získání vědecké hodnosti či jiné odborné kvalifikace vyššího stupně.
 - o **Habilitační práce** je původní vědecká práce, za niž nebyla získaná žádná jiná akademická nebo vědecká hodnost. Je podmínkou k získání vědecko-pedagogického titulu docent (Matušík 2003).
 - o **Disertační práce** je dokument, jenž prezentuje autorův výzkum a jeho výsledky a je autorem předkládán k řízení pro získání vědecké hodnosti či jiné odborné kvalifikace vyššího stupně (Matušík 2003).
 - o **Rigorózní práce** prokazuje schopnost samostatné činnosti v oblasti výzkumu a vývoje nebo k samostatné tvůrčí činnosti. Rigorózní práce, resp. její tvorba a následná obhajoba je součástí tzv. státní rigorózní

zkoušky, kterou musí student složit, aby mu mohl být udělen tzv. malý doktorát. Jmenovitě se může jednat o tituly například JUDr., PhDr., RNDr., ThDr. nebo ThLic. Často vychází či navazuje na úspěšně obhájenou a dobře hodnocenou diplomovou práci s tím, že je kladen důraz na práci se zahraniční literaturou a také se očekává od autora rigorózní práce kvalitnější výzkumné šetření. Rigorózní práce je v podstatě mezistupeň mezi pregraduálním a postgraduálním stupněm vysokoškolského studia.

- o **Diplomová práce** je dokument vyžadovaný ke státní závěrečné zkoušce na vysoké škole, popř. k závěrečné zkoušce na střední odborné škole. Obsahuje popis a výsledky samostatné odborné práce na předem zadané téma. Její předložení a obhájení je součástí řízení pro přiznání akademického titulu (Bc., Mgr., Ing., MUDr. a MVDr.) (Matušík 2003).
- o **Bakalářská práce** je dokument vyžadovaný ke státní závěrečné zkoušce na vysoké škole, popř. k závěrečné zkoušce na střední odborné škole. Obsahuje popis a výsledky samostatné odborné práce na předem zadané téma. Její předložení a obhájení je součástí řízení pro přiznání akademického titulu bakalář (Bc.) (Matušík 2003).
- **Zprávy** - nová, souhrnná či cílená jednorázová informace o projektech, výzkumných záměrech, nebo institucích.
 - o **Výroční zpráva** je oficiální dokument instituce, jehož vyhotovení ustanovuje Zákon č. 563/1991 Sb. v §21. Účelem výroční zprávy je uceleně, vyváženě a komplexně informovat o vývoji, výkonnosti, činnosti a stávajícím hospodářském postavení instituce, většinou formou brožury obsahující zákonem dané finanční údaje a další důležité informace. Výroční zpráva obsahuje též účetní závěrku a zprávu auditora.
 - o **Výzkumná zpráva** je dokument, který souvisí s přípravou, realizací a hodnocením určitého vědeckého výzkumu a který obsahuje rozbor i charakteristiku průběhu a výsledků řešení daného výzkumného úkolu.
 - o **Závěrečná zpráva** z projektu je zpráva uvádějící údaje o řešení a konečných výsledcích projektu. Zpráva se zpravidla skládá ze dvou částí. První částí je popis průběhu řešení projektu a popis dosažených výsledků. Druhou částí pak je popis vyúčtování projektu.
 - o **Průběžná zpráva** z projektu je zpráva uvádějící údaje o řešení, průběhu a dílčích výsledcích projektu. Zpráva se zpravidla skládá ze dvou částí. První částí je popis průběhu řešení projektu a popis dosažených výsledků a druhou částí pak je popis vyúčtování projektu.
 - o **Technická zpráva** je dokument související s přípravou, realizací a hodnocením určitého výzkumu, který obsahuje rozbor a charakteristiku průběhu i výsledků řešení daného úkolu. Je výsledkem činnosti průmyslových podniků, komerčních organizací, vědeckých institucí apod.
 - o **Grantová zpráva** je zpráva uvádějící údaje o řešení grantu, jeho postupu a dosažených výsledcích. Součástí grantové zprávy bývá finanční výkaz, účetní uzávěrka, informace o výsledcích a dopadech, plány udržitelnosti aj.
 - o **Zpráva z průzkumu** je zpráva uvádějící údaje o řešení průzkumu, jeho postupu a dosažených výsledcích. Součástí zpravidla bývá plné znění odpovědí respondentů.
 - o **Statistická zpráva** -je zpráva obsahující shromážděné, analyzované a zpracované informace, které kvantitativně charakterizují zákonitosti jevů ve spojitosti s jejich kvalitativním obsahem.

- **Konferenční materiály** - dokumenty vznikající v souvislosti s konáním konferencí.
 - o **Sborník** je publikace obsahující dva či více samostatných textů, zpravidla alespoň rámcově tematicky příbuzných, se společným názvem. Může být koncipován jako publikace jednorázová nebo seriálová (Matušík 2003).
 - o **Příspěvek z konference** je prezentace nebo plný text příspěvků v souladu s tématem konference oficiálně přednesených na konferenci a určena k publikaci v konferenčním sborníku.
 - o **Poster** je velký, zpravidla jednobližňově výtvarně zpracovaný grafický dokument s tiskem po jedné straně (někdy s volným místem pro doplnění informace) informačního, reklamního nebo ideologicky apelativního charakteru určené k prezentaci na veřejném místě (Matušík 2003).
- **Autorské práce** - dokumenty, které jsou výsledkem intelektuální činnosti jednotlivých osob.
 - o **Preprint** je zpravidla nazývána verze článku před recenzním řízením.
 - o **Referát** je redukovaný text, který bez doplňkových nebo hodnotících informací charakterizuje obsah dokumentu. Základními vlastnostmi referátu jsou výstižnost, přehlednost, jasnost, stručnost, přesnost, objektivnost a čtivost. Referát je formulován v přirozeném jazyce - obvykle ve větách, někdy však jen v heslovité podobě, telegrafickým stylem. Referát může používat textových formulací z referovaného dokumentu, ale jako celek je formulován nově. Ve spojení s bibliografickými záznamy referovaných dokumentů se referáty publikují v referátových publikacích (např. referátové bibliografie, referátová periodika apod.) Podle charakteru zpracování lze vydělit informativní, indikativní, kritický, modulární, analytický, monografický, přehledový, výběrový a autorský referát (Švejda 2003)
- **Studijní materiály** - dokumenty vznikající při výuce na jednotlivých typech vysokých škol, jejichž autory mohou být pedagogičtí pracovníci nebo studenti.
 - o **Učební osnovy** jsou dokumenty obsahující detailní informace k probíhající výuce v jednotlivých předmětech.
 - o **Témata** ke zkouškám jsou dokumenty obsahující buď konkrétní otázky, nebo okruhy témat ke zkouškám. Jsou vypracovány vyučujícím pro studenty.
 - o **Záznamy z přednášek** jsou vyučujícím nebo studentem zpracované zápisy z přednášek včetně např. prezentací.
- **Firemní literatura** - dokumenty vznikající v komerčních i nekomerčních organizacích v rámci pracovního procesu a informující veřejnost o produktech, poskytovaných službách a činnostech organizace.
 - o **Katalog výrobků** je oficiální dokument vydávaný podnikatelskými subjekty za účelem informování o nabídce zboží či služeb.
 - o **Manuál** je podrobný návod (manuál) k nějaké činnosti, výrobě, použití nebo opravě (Matušík 2003).

V roce 2011 vznikla první verze mezinárodní typologie šedé literatury, <http://code.google.com/p/grey-literature-typology/issues/list>, kterou doporučujeme využít.

3.2.2 Obsahové hledisko

Z obsahového hlediska je profil digitálního repozitáře vymezen podle zaměření konkrétní instituce, kterou je repozitář budován. Institucím doporučujeme přihlídnout k typologiím a obsahovým vymezením národních, resp. mezinárodních repozitářů, se kterými chtějí do budoucna spolupracovat.

3.2.3 Časové hledisko

Určení časového hlediska (tj. roky sběru digitálních dokumentů) je důležité z několika důvodů. Z pohledu práva je nutné aplikovat vždy ten výklad práva, který byl platný v roce vzniku dokumentu. Např. zaměstnanecká díla má právo vkládat, tj. má autorská práva majetková, instituce od roku 2000 výše, ale pokud se jedná o zaměstnanecké dílo vytvořené před rokem 2000, musí instituce požádat o souhlas autora, byť byl zaměstnancem instituce (viz kapitola 3.1.1 Autorský zákon). Z pohledu zpracování je potřeba připravit dokument vložením do repozitáře různým způsobem. Např. tištěné dokumenty je potřeba nejprve digitalizovat nebo digitální dokumenty z 80. či 90. let je potřeba převést do současných formátů apod.

Doporučujeme neomezovat vkládání dokumentů pouze na poslední roky, kdy jsou dokumenty zpracovány většinou tak, že je není nutné předem upravovat, jelikož obsahová kvalita starších dokumentů bývá vysoká a je přínosné ji do repozitáře vkládat i přes náročnost jejich přípravy.

3.2.4 Forma

Z hlediska formy dokumentu lze do digitálního repozitáře zařadit digitální dokumenty, a to jak born-digital⁸ dokumenty, tak digitalizované dokumenty. Pro správu je vhodné si předem určit, které druhy digitálních forem bude repozitář podporovat (např. text, obraz, multimédia, prezentace, strukturované dokumenty apod.). U každého druhu digitální formy je pak důležité zvolit podporované formáty pro dlouhodobé uložení a zpřístupnění. Problematika dlouhodobého uložení a zpřístupnění bude podrobně rozebrána v kapitole 5.1.2 Formáty dokumentů.

3.2.5 Geografické hledisko

Instituce si určí, zda bude zařazení dokumentů vkládaných do digitálního repozitáře omezovat místem vzniku dokumentu (např. na území České republiky). Toto hledisko má význam především u institucí s pobočkami nebo v případě budování mezinárodních repozitářů. Při budování institucionálních repozitářů s jedním místem působnosti není geografické hledisko zásadním ukazatelem.

3.2.6 Stupeň zveřejnění

Instituce si určí, zda do digitálního repozitáře bude zařazovat pouze polopublikované (např. příspěvky ze sborníku) a nepublikované dokumenty (např. výzkumné zprávy), tedy pouze šedou literaturu. Anebo zda bude do digitálního repozitáře vkládat také publikovanou literaturu. V takovém případě může instituce vkládat publikační činnost svých zaměstnanců, pokud to není zapovězeno v licenční smlouvě s vydavatelem.

⁸ Termín „born-digital“ odkazuje na materiály, které vznikají v digitální podobě. A to v kontrastu s digitálním přeformátováním, jejímž prostřednictvím se analogové materiály stávají digitálními. Tento případ je nejčastěji používán ve vztahu k digitálním knihovnám a problematice organizace uchování digitálních záznamů a duševního vlastnictví. Nicméně, nezávisle na tom jak je vývoj technologií rychlý a pokročilý, je koncepce „born-digital“ diskutována především díky nástupu e-knih a rozvíjející se digitální hudbě v souvislosti s osobními potřebami spotřebitelů. Termíny, které by mohly být považovány za synonymní jsou „natively digital“, „digital-first“, a „digital-exclusive“. *Digital-born* [online]. Wikipedia, 2011. [cit. 2011-08-03]. Dostupné z [www: http://en.wikipedia.org/wiki/Born-digital](http://en.wikipedia.org/wiki/Born-digital)

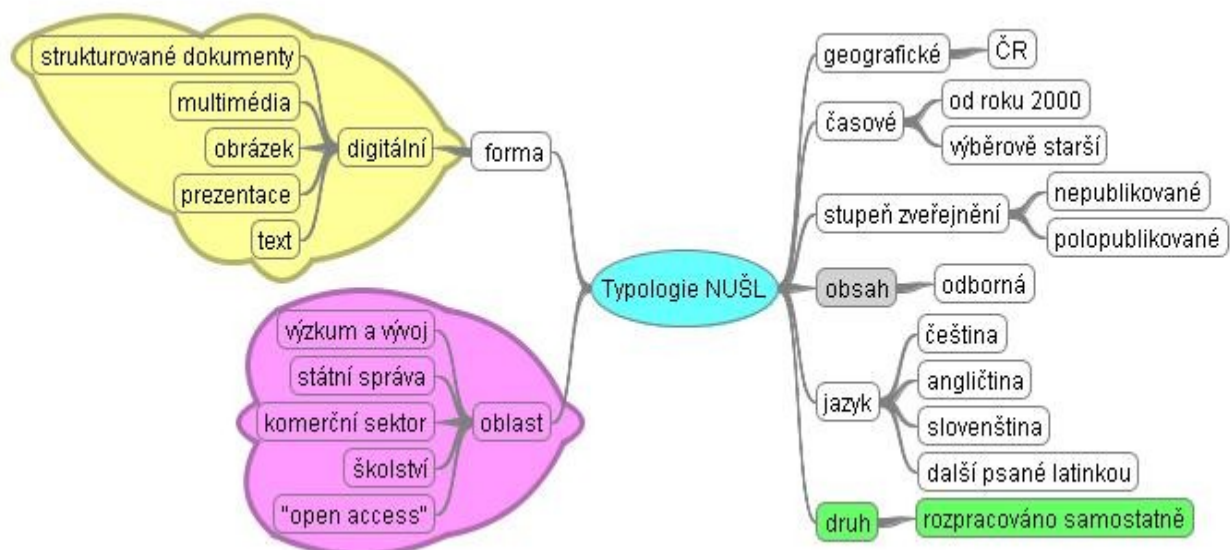
Možnost zařazení publikované literatury do digitálního repozitáře je zde uvedena z důvodu komplexního pohledu na určení obsahu digitálního repozitáře. Je nutné upozornit na to, že tato metodika se netýká publikované literatury.

3.2.7 Jazykové hledisko

Instituce si definuje výčet jazyků, ve kterých budou dokumenty ukládané v digitálním repozitáři. Toto hledisko je důležité pro správce systému především v případě, že budou vybrány jazyky v jiných znakových sadách, než je latinka. Doporučená znaková sada pro ukládání dat např. v databázi repozitáře a kódování stránek je UTF-8.

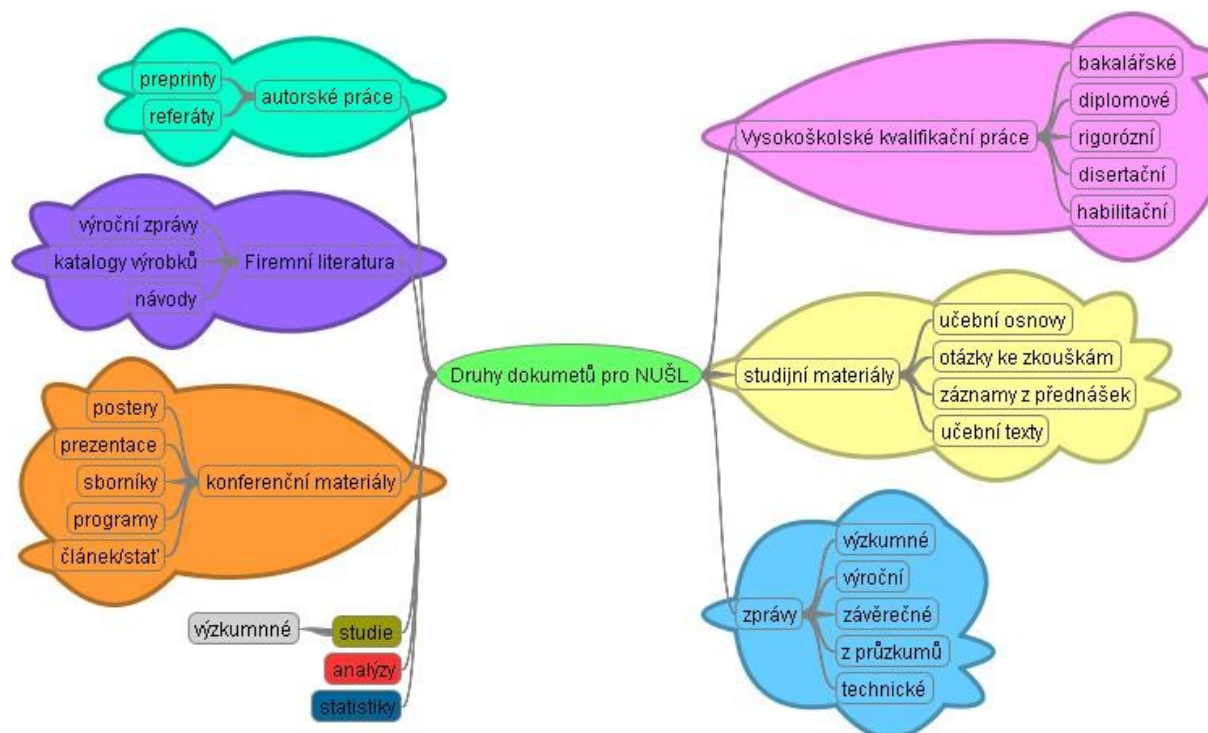
Příklad NUŠL:

Příklad zpracování všech zmíněných hledisek typologie dokumentů NUŠL, na jejímž základě je určen obsah repozitáře. Jednotlivá hlediska jsou znázorněna v následujícím diagramu.



Obrázek 1: Základní typologie NUŠL

Ze základní typologie dokumentů je rozpracována charakteristika druhu dokumentu, kde jsou hlavními skupinami vysokoškolské kvalifikační práce, studijní materiály, zprávy, autorské práce, firemní literatura, konferenční materiály a další. Jednotlivé skupiny již obsahují konkrétní druhy dokumentů, které jsou znázorněny v následujícím diagramu.



Obrázek 2: Typologie dle druhu dokumentu

3.3 Výběr software pro digitální repozitář

Výběr softwarového řešení pro digitální repozitář se v zásadě neliší od výběru softwaru pro jakoukoliv jinou činnost. (PEJŠOVÁ, 2010b, s. 93 – 106.) Jakkoliv se může z digitální povahy úložiště zdát, že softwarový systém je hlavní součástí plánovaného repozitáře, je to nakonec pouze nástroj, který umožní repozitář realizovat. Nejdůležitější přípravnou fází pro výběr softwaru je tedy dobrý popis samotného repozitáře, jeho cílů, hlavních funkcí, procesů, činností, obsahu a vazeb na okolí. Z dobré definice repozitáře pak lze snadno odvodit hlavní požadavky na softwarový systém a kritéria jeho výběru.

Hlavní kritéria výběru software:

- 1) řízení obsahu digitální knihovny,
- 2) uživatelské rozhraní,
- 3) správa a řízení přístupu uživatelů,
- 4) administrace systému,
- 5) jiné požadavky.

3.3.1 Řízení obsahu digitálního repozitáře

Toto kritérium se vztahuje k hlavním funkcím digitálního repozitáře. Zahnuje tvorbu, ukládání, kontrolu a verzování obsahu (např. při migraci souborů) digitálního repozitáře, jeho prohlížení lineární (rejstříky) i hierarchické, správu metadat, různé mechanismy pro získávání dat (např. harvesting - OAI-PMH, web services), interoperabilitu pro spolupráci s jinými digitálními repozitáři (povinně OAI-PMH, dále např. SRU), podporu pro velké objemy dat, vícejazyčné vyhledávání nad metadaty a soubory (fulltextové), logické a proximní vyhledávání, podporu různých formátů

ukládání souborů - textových (např. Adobe PDF, MS DOC, MS DOCx, XML (HTML), TXT), obrazových (např. JPG 2, JPG, PNG, TIFF), prezentačních (např. Adobe PDF, MS PPT, MS PPTx), strukturovaných (např. HTML a XML), audio (např. MP3, AVI) a video (např. MP4, MPEG-1, MPEG-2, AVI), podpora kódování (př. ASCII, UNICODE, UTF-8).

3.3.2 Uživatelské rozhraní

Přístup přes www s podporou obecně rozšířených internetových prohlížečů, možnost přizpůsobení uživatelského rozhraní pro různé role uživatelů zvenku i zevnitř instituce, v případě potřeby vícejazyčné uživatelské rozhraní a vícejazyčná prezentace výsledků vyhledávání. Výběr software musí zohlednit i kvalitu uživatelského rozhraní. V tomto směru je nutné sledovat trendy vývoje pro komunikaci člověk – počítač.

3.3.3 Správa a řízení přístupu uživatelů

Klasifikace a profily uživatelů na základě jejich rolí (např. RBAC), autentifikace uživatelů na základě uživatelského jména a hesla, řízení přístupu podle IP adres, přístup přes proxy, monitorování a reporting pro vyhodnocování chování uživatelů pro pozdější využití ke zlepšení služeb digitálního repozitáře, autorizace uživatelů (např. LDAP, Shiboleth, Eduroam, OpenID), zaznamenávání aktivit uživatelů pro účely účtování.

3.3.4 Administrace systému

Tento požadavek se vztahuje ke správě provozního prostředí digitálního repozitáře, která může být zejména u digitálních knihoven velkého rozsahu velice pracná a náročná. Spadají sem nástroje pro správu metadat, nastavování automatických mechanismů sběru dat, indexování, automatické generování klíčových slov, formátování výstupů apod.

3.3.5 Jiné požadavky

Uživatelská a administrátorská dokumentace, helpdesk (podpora ze strany výrobce/implementátora systému), rozšiřující funkce pro uživatele jako jsou diskusní a komunikační skupiny, automatické upozorňování na různé události (výskyt nových dokumentů splňující určitá kritéria), RSS výstupy, uživatelské schránky a komentáře k dokumentům, požadavky na provozní hardware a software apod. (PEJŠOVÁ, 2010b, s. 93 – 106.)

Doporučení:

Určitým specifickým rysem v oblasti softwaru pro digitální repozitáře, jež se provozují převážně v neziskovém sektoru, je například v porovnání s komerční sférou dostupnost programového vybavení z oblasti Open Source (software s otevřeným zdrojovým kódem, dostupný zpravidla bez licenčních poplatků), jehož několik nejznámějších představitelů (např. CDS Invenio, DSpace, Eprints, Fedora a Greenstone) má za sebou úspěšné instalace, z nichž se lze poučit.

4 Zpracování šedé literatury

Zpracování dokumentu

Pro zpracování záznamů o šedé literatuře je možné zvolit již existující nebo definovat vlastní metadatový formát, v jakém se budou údaje o těchto dokumentech zpracovávat. Výhodou definování vlastního formátu je možnost zohlednit vlastní specifika zpracovávaných dokumentů a přístup k popisu. Nevýhodou vlastního formátu je jednak nekompabilita v okolí a možnost opakování chyb, a také intelektuální i časová náročnost definování vlastního formátu, potenciální problémy při předávání dat do národních a mezinárodních registrů. Výhodou využití již existujícího a tedy ověřeného formátu je ušetření pracovní kapacity zaměstnanců, prověřenost takového formátu v praxi a zkušenosti s odevzdáváním takto zpracovaných dat do jiných systémů. Nevýhodou je potřeba přizpůsobit popis dokumentu zvolenému formátu.

Pro zápis dat ve vybraném formátu je vhodné použít šablony pro vkládání jednotlivých typů dokumentů. Zápis do šablon zpřehlední zpracování metadatového záznamu, protože nabízí pouze prvky pro vybraný typ dokumentu a není tak nutné procházet plný výčet prvků v celé bohatosti formátu. Smyslem této části je poskytnout instrukce, v jakém formátu a jakým způsobem zpracovávat údaje o šedé literatuře.

Při zpracování záznamů do digitálního repozitáře je třeba zohlednit:

- 1) metadatový formát a specifické prvky pro šedou literaturu,
- 2) šablony pro jednotlivé druhy dokumentu,
- 3) vkládání dat do digitálního repozitáře.

4.1 Metadatový formát

Formát pro uložení metadat je obecně nezbytnou součástí budování repozitářů. Je tedy neopomenutelný též při budování repozitářů šedé literatury. Metadata charakterizuje Česká terminologická databáze knihovnictví a informační vědy (dále jen TDKIV) takto:

„Strukturovaná data, která nesou informace o primárních datech. Pojem metadat je používán především v souvislosti s elektronickými zdroji a vztahuje se k datům v nejširším smyslu slova (datové soubory, textové informace, obrazové informace, hudba aj.). Funkce metadat je popisná, selekční a archivační. V souvislosti s těmito funkcemi se rozlišují metadata pro účely popisu, správy, právních nároků, technické funkčnosti, užití a archivace. Údaje se obvykle vkládají přímo do zdroje (umísťují se např. v záhlaví dokumentu HTML).“ (KTD, 2003)

4.1.1 Formát NUŠL

Metadatový formát NUŠL byl speciálně vytvořený pro zpracování záznamů o digitálních dokumentech šedé literatury. Tento formát je lineární, neumožňuje tedy analytický popis např. propojení sborníku a jeho příspěvků.

Metadatový formát NUŠL používá prvky z formátů Dublin Core Metadata Element Set, DCMI Metadata Terms, EVSKP-MS, EDT-MS a vlastní prvky.

Formát NUŠL je otevřený formát, který může být v budoucnu doplňován o elementy, které budou nutné pro provoz digitální knihovny šedé literatury a spolupráci jak s lokálními dodavateli dat, tak s mezinárodními registry. V lokálních repozitářích je možné využít dalších prvků nad rámec formátu NUŠL podle potřeb instituce.

Formát NUŠL se skládá z povinných a nepovinných prvků. Podrobný popis celého formátu i s příklady u jednotlivých prvků je dostupný na webových stránkách projektu NUŠL http://nusl.techlib.cz/images/NUSLformat_verze1.0_1.pdf

Popisná metadata

Popisná metadata zachycují informace o samotném digitálním zdroji a slouží k jeho vyhledání v digitálním repozitáři.

4.1.1.1 Povinné prvky NUŠL

nusl:metadata

Kořenový prvek metadatového záznamu.

dc:identifier

Tento prvek je do formátu NUŠL převzatý z Dublin Core (dc:identifier). Jednoznačný a trvalý identifikátor zdroje.

dc:title

Tento prvek je do formátu NUŠL převzatý z Dublin Core. Hlavní název zdroje, který zdroji přidělil tvůrce.

dc:title (nusl:typeTranslated)

Tento prvek je do formátu NUŠL převzatý z Dublin Core. Hlavní název zdroje v anglickém jazyce, pokud není uveden v poli dc:title.

dc:creator

Tento prvek je do formátu NUŠL převzatý z Dublin Core. Fyzická osoba, která je primárně odpovědná za vytvoření obsahu zdroje, tj. autor VŠKP, hlavní řešitel projektu, vedoucí výzkumného záměru, apod., příp. korporace, která nese primární autorskou odpovědnost za dílo.

dcterms:dateCreated

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms. Datum zveřejnění zdroje.

dc:subject

Tento prvek je do formátu NUŠL převzatý z Dublin Core. Obsahem tohoto prvku je věcný popis šedé literatury vyjádřený buď ve formě termínů z Polytematického strukturovaného hesláře (dále jen PSH), nebo ve formě volně tvořených klíčových slov, a to v českém nebo anglickém jazyce.

dc:type

Tento prvek je do formátu NUŠL převzatý z Dublin Core. Obsahem je typologie dokumentu NUŠL dle druhu dokumentu.

dc:language

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Jazyk intelektuálního obsahu zdroje.

dc:rights

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Textová informace o právech k využívání. Specifikují se podmínky, na základě kterých může být práce distribuována, reprodukována atd.

4.1.1.2 Nepovinné prvky NUŠL

dcterms:alternative

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms.
Podnázev zdroje nebo další informace k hlavnímu názvu.

dc:contributor

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Spoluautor dokumentu – fyzická osoba, která významně intelektuálně přispěla k vytvoření zdroje, ale jejíž přínos je sekundární ve srovnání s osobami či entitami uvedenými v prvku Autor (např. editoři, přepisovači, ilustrátoři a pořadatelé).

dc:publisher

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Instituce odpovědná za uložení, trvalou archivaci nebo zpřístupnění práce v elektronické formě. Může jít o instituci jako celek nebo o její podřízenou jednotku.

dcterms:dateModified

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms.
Datum a popř. čas ukončení procesu modifikace zdroje (například v rámci migrace souborů za účelem trvalého uložení práce v digitálním úložišti, popř. datum vzniku elektronické verze zdroje pokud se jedná o digitalizovaný dokument). Údaje lze zjistit z vlastností souboru.

nusl:note

Nově vytvořený prvek pro NUŠL.
Poznámka k dokumentu, údaj k digitálnímu zdroji, který nelze uvést v jiném poli.

dc:subject

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Opakovaný výskyt, téma obsahu zdroje.

dcterms:abstract

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms.
Textový popis zahrnující detaily a vztahující se k pojednávanému tématu v českém nebo anglickém jazyce.

dcterms:bibliographicCitation

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms.
Stručný záznam ve formě bibliografické citace. Zapisuje se v souladu s mezinárodním standardem ČSN ISO 690 a ČSN ISO 690-2 (s ohledem na české prostředí).

dc:rights

Tento prvek je do formátu NUŠL převzatý z Dublin Core.
Opakovaný výskyt, je možné doplnit některou z CC licencí.

dcterms:extent

Tento prvek je do formátu NUŠL převzatý z Dublin Core Terms.
Rozsah nebo velikost práce. V případě potřeby se zapisuje přesný rozsah práce (počet stran, příloh apod.), délka trvání zvukového záznamu aj.

4.1.1.3 Specifické prvky pro šedou literaturu

Dokumenty šedé literatury obsahují některé specifické údaje. Pro zápis těchto údajů o dokumentu byly definovány nové prvky, protože odpovídající elementy nebyly součástí žádného z analyzovaných metadatových formátů. Vyjma skupiny prvků pro eVŠKP a studijní materiály, které byly převzaty z formátu EVŠKP-MS. Tyto prvky jsou též součástí formátu NUŠL, ale byly vyčleněny do samostatné části, protože je velmi důležité uvědomit si významnost těchto prvků při budování institucionálních repozitářů. Přes jejich důležitost nejsou zařazeny mezi povinné prvky, protože se vztahují vždy jen ke konkrétním druhům šedé literatury.

Elementy pro zpracování zpráv

Elementy týkající se zpracování zpráv jsou pro NUŠL nově vytvořené a inspirací pro jejich vznik byl především Automatizovaný systém evidence publikací (dále jen ASEP)⁹.

nusl:resultID

Nově vytvořený prvek pro NUŠL.

Obsahem opakovatelného pole je označení grantu (grantů), v jehož rámci byla vypracována hlášená publikace. Obsahem pole může být také označení výzkumného záměru pracoviště, v jehož rámci byla vypracována hlášená publikace.

nusl:resultGrantGarant

Nově vytvořený prvek pro NUŠL.

Obsahem prvku je název poskytovatele projektu.

Elementy pro zpracování konferenčních materiálů

Oblast popisu konferenčních materiálů nebyla v žádném z formátů, které autoři NUŠL formátu považovali za zdrojové nebo výchozí pro tvorbu formátu NUŠL podrobněji zpracována. Rozhodli proto pro vytvoření vlastních prvků.

Prvky tak, jak jsou nyní používány, jsou také vhodné pro spolupráci se systémem OpenGrey¹⁰, protože v podstatě odpovídají tomu, jak jsou v tomto systému zpracovávány konferenční materiály. S tímto systémem byla rovněž sladěna pravidla pro vyplnění jednotlivých polí.

nusl:event

Nový element vytvořený pro formát NUŠL.

⁹ Systém ASEP obsahuje bibliografické záznamy o výsledcích vědeckého výzkumu ústavech AV ČR od roku 1985, v úplnosti od roku 1993. Záznamy ze systému ASEP se používají také pro zasílání do databáze RIV.

¹⁰ Evropská databáze šedé literatury z oblasti vědy, techniky, ekonomie, humanitních a dalších věd. Více v *kapitole 5.4.6 Mezinárodní agregátory zaměřené šedé literatury*

Název konference nebo jiné akce. Oficiální název konference nebo akce se zapisuje v originálním jazyce, v nezkrácené podobě.

nusl:eventPlace

Nový element vytvořený pro formát NUŠL.

Stát (místo) konání konference nebo jiné akce. Zapisuje se ve formě město, v závorkách uvedený stát konání konference.

nusl:eventDates

Nový element vytvořený pro formát NUŠL.

Datum nebo rozmezí dat konání konference nebo jiné akce, zápis dle normy ISO 8101.

Elementy pro zpracování eVŠKP

Následující skupina prvků pro eVŠKP a studijní materiály byla převzata z formátu EVŠKP-MS, resp. ETD-MS: an Interoperability Metadata Standard for Electronic Theses and Dissertations. Prvky se týkají údajů o akademickém titulu nebo vědecko-pedagogické hodnosti a instituci, která ho přiděluje, a jsou vnořené do prvku thesis:degree.

thesis:name

Tento prvek je do formátu NUŠL převzatý z formátu EVŠKP-MS.

Zkratka jména akademického titulu nebo vědecko-pedagogické hodnosti přidělované vysokou školou (například Ph.D., doc., JUDr., Mgr., Ing., aj.).

Údaje o akademickém titulu nebo vědecko-pedagogické hodnosti získané v souvislosti se zpracovávaným dokumentem.

thesis:level

Tento prvek je do formátu NUŠL převzatý z formátu EVŠKP-MS.

Typ studijního programu, v jehož rámci byla připravena VŠKP (doktorský, magisterský, bakalářský aj.).

thesis:discipline

Tento prvek je do formátu NUŠL převzatý z formátu EVŠKP-MS.

Studijní program a studijní obor akreditovaný příslušnou komisí Ministerstva školství, mládeže a tělovýchovy (dále jen MŠMT), v jejichž rámci byla připravena a obhájena VŠKP.

thesis:grantor

Tento prvek je do formátu NUŠL převzatý z formátu EVŠKP-MS.

Instituce (škola, ústav apod.) garantující přidělení titulu spojené s vytvořením VŠKP.

4.1.1.4 Technická a administrativní metadata

Technická a administrativní metadata jsou využívána při předávání záznamů, popř. digitálních zdrojů do dalších databází a registrů.

nusl:accessRights

Nově vytvořený prvek pro NUŠL.

Informace o zveřejnitelnosti dokumentu.

nusl:provider

Nově vytvořený prvek pro NUŠL.
Název instituce, která je odpovědná za záznam.

nusl:fileType

Nově vytvořený atribut pro NUŠL.
Popis digitálního dokumentu.

nusl:saveRecord

Nově vytvořený prvek pro NUŠL.
Datum uložení záznamu do systému.

nusl:accessibility

Nově vytvořený prvek pro NUŠL.
Informace o dostupnosti dokumentu v případě, že digitální dokument není k dispozici.

Další informace k formátu NUŠL naleznete na stránkách projektu NUŠL
<http://nusl.techlib.cz/index.php/Metadata> .

4.1.2 Formát EVSKP- MS

Formát EVSKP-MS definuje národní soubor prvků pro popis VŠKP obhájovaných na vysokých školách v České republice a pro přenos souborů.¹¹ Základní skupina prvků tohoto formátu vychází z formátu Dublin Core, některé specifické prvky pro popis VŠKP byly převzaty z formátu ETD-MS. Z koncepčních i jiných hledisek pak byly další inspirací některé národní registry, např. XMetadiss a TEF 2.0, zejména pak definice jednotlivých prvků. Podrobný popis prvků pak přihlíží k praxi popisu VŠKP v ČR.

Formát EVSKP-MS je doporučeným výměnným formátem pro repozitáře VŠKP v ČR.

4.1.2.1 Povinné prvky EVSKP-MS

evskp:metadata

Kořenový prvek metadatového záznamu.

dc:title

Hlavní název VŠKP.

dcterms:abstract

Textový popis, zahrnující detaily vztahující se k tématu pojednávanému ve VŠKP, zejména cíle práce, hypotézy, metody, závěry apod.

dc:creator

Fyzická osoba, která je primárně odpovědná za vytvoření obsahu VŠKP, tj. její autor.

dcterms:dateAccepted

Přesné datum obhajoby VŠKP uvedené v protokolu obhajoby.

dc:type

Typ VŠKP.

dcterms:medium

¹¹ <http://www.evskp.cz/standardy/evskp/>

Označení datového formátu digitálního souboru hlavní části VŠKP, ve kterém je uložena a zpřístupňována.

dc:identifier

Jednoznačný identifikátor VŠKP. Může jít o trvalý identifikátor nebo může jít o označení místa uložení VŠKP v elektronické formě, z něhož ji uživatel může získat (například URL).

dc:language

Jazyk intelektuálního obsahu VŠKP. Jde o hlavní jazyk, v němž je VŠKP napsána.

thesis:degree

Údaje o akademickém titulu nebo vědecko-pedagogické hodnosti a instituci, která ho přiděluje.

thesis:name (prvek vnořený v prvku thesis:degree)

Zkratka jména akademického titulu nebo vědecko-pedagogické hodnosti přidělované vysokou školou (například Ph.D., doc., JUDr. Mgr., Ing. aj.).

thesis:level (prvek vnořený v prvku thesis:degree)

Typ studijního programu, v jehož rámci byla připravena VŠKP (doktorský, magisterský, bakalářský aj.).

thesis:discipline (prvek vnořený v prvku thesis:degree)

Studijní program a studijní obor akreditovaný příslušnou komisí MŠMT, v jejichž rámci byla připravena a obhájena VŠKP.

thesis:grantor (prvek vnořený v prvku thesis:degree)

Instituce (škola, ústav apod.) garantující přidělení titulu spojené s vytvořením VŠKP.

4.1.2.2 Nepovinné prvky EVSKP-MS

dcterms:alternative

Podnázev VŠKP nebo další informace k hlavnímu názvu.

dc:subject

Téma VŠKP vyjádřené pomocí volně tvořených klíčových slov, nebo termíny z některého z řízených slovníků, nebo třídníků klasifikačních systémů.

dcterms:tableOfContents

Obsah VŠKP, např. seznam kapitol.

dc:publisher

Instituce odpovědná a uložení, trvalou archivaci a nebo zpřístupnění VŠKP v elektronické formě.

dc:contributor

Fyzická osoba odpovědná za vedení nebo oponování VŠKP.

dcterms:created

Datum vytvoření VŠKP.

dcterms:dateSubmitted

Datum odevzdání, provedení či podání VŠKP.

dcterms:modified

Datum a popř. čas ukončení procesu modifikace VŠKP (například v rámci migrace souborů za účelem trvalého uložení práce v digitálním úložišti).

dcterms:extent

Rozsah nebo velikost VŠKP.

dcterms:bibliographicCitation

Stručný záznam VŠKP ve formě bibliografické citace.

dc:rights

Textová informace o právech k využívání VŠKP. Specifikují se podmínky, na základě kterých může být práce distribuována, reprodukována atd.; stanovit lze i dobu, po kterou dané podmínky platí a kontakt na osoby, které se o práva starají.

4.1.2.3 Technická a administrativní metadata EVSKP-MS

evskp:contact

Identifikátor poskytovatele metadat.

evskp:fileNumber

Počet souborů VŠKP.

evskp:fileProperties

Popis konkrétního souboru VŠKP.

evskp:transfer

Identifikátor odkazující na soubor tvořící VŠKP nebo archiv ZIP se soubory obsahujícími VŠKP.

evskp:server

Informace o serveru zpřístupňujícím VŠKP.

evskp:dateDelivered

Datum doručení metadatového záznamu do repozitáře.

dcterms:available

Zpřístupnění souborů VŠKP.

evskp:modified

Datum změny záznamu VŠKP.

Další informace k formátu EVŠK-MS naleznete na stránkách <http://www.evskp.cz/standardy/evskp/>.

Doporučení:

Při budování digitálního repozitáře doporučujeme využívat stávající metadatové formáty a nezabývat se vývojem vlastních nových formátů. Pro zpracování záznamu výhradně o VŠKP doporučujeme využít formát EVSKP-MS, pro ostatní druhy šedé

literatury pak formát NUŠL. Pokud je potřeba interně uložit metadata v institucionálním repozitáři v jiném než doporučeném formátu, je nutné minimálně ukládat odpovídající povinné metadatové prvky podle doporučených formátů z důvodu možnosti konverze dat při převodu do národních systémů NUŠL a Thesis.cz. Spolupráce s těmito systémy je popsána v kapitole 5.4.5 Národní agregátory zaměřené šedé literatury. Pokud institucionální repozitář odevzdává data do repozitáře NUŠL ve formátu NUŠL nebo VSKP-MS, je potřebná konverze a zasílání metadat do databáze OpenGrey zajištěno prostřednictvím repozitáře NUŠL.

4.2 Šablony pro šedou literaturu

Při budování digitálního repozitáře NUŠL byly vytvořeny šablony (formuláře) pro vkládání jednotlivých druhů dokumentu. Tyto šablony jsou použitelné pro jakýkoliv digitální repozitář šedé literatury. Šablony jsou vytvořeny na základě povinných a nepovinných prvků metadatového formátu NUŠL pro jednotlivé druhy dokumentů šedé literatury.

Součástí každé šablony je popis skupiny dokumentů, pro které je šablona připravena, charakteristika jednotlivých dokumentů a výpis povinných polí pro danou skupinu dokumentů.

4.2.1 Šablona pro zpracování ZPRÁV

Zprávy - nová, souhrnná či cílená jednorázová informace o projektech, výzkumných záměrech, nebo institucích.

Tato šablona je připravena pro zpracovávání těchto typů zpráv:

- výroční zprávy
- průběžné zprávy z projektů
- závěrečné zprávy z projektů
- grantové zprávy
- výzkumné zprávy
- technické zprávy
- statistické zprávy

Seznam povinných polí:

- Identifikátor (dc:identifier)
- Název (dc:title)
- Autor (dc:creator)
- Datum zveřejnění zdroje (dcterms:dateCreated)
- Předmět (dc:subject)
- NUŠL typ dokumentu (dc:type)
- Jazyk dokumentu (dc:language)
- Autorská práva (dc:rights)
- Omezení přístupu (nusl:transfer)
- Název instituce (nusl:provider)
- Popis souboru (nusl:filetype)

4.2.2 Šablona pro zpracování KONFERENČNÍCH MATERIÁLŮ

Konferenční materiály - dokumenty vznikající v souvislosti s konáním konferencí.

Tato šablona je připravena pro zpracovávání těchto typů konferenčních materiálů:

poster
příspěvky z konference
sborníky
programy

Seznam povinných polí:

Identifikátor (dc:identifier)
Název (dc:title)
Autor (dc:creator)
Název konference/akce (nusi:event)
Místo konání konference/akce (nusi:eventPlace)
Datum nebo rozmezí konání konference/akce (nusi:eventDates)
Datum zveřejnění zdroje (dcterms:dateCreated)
Předmět (dc:subject)
NUŠL typ dokumentu (dc:type)
Jazyk dokumentu (dc:language)
Autorská práva (dc:rights)
Omezení přístupu (nusi:transfer)
Název instituce (nusi:provider)
Popis souboru (nusi:filetype)

4.2.3 Šablona pro zpracování vysokoškolských KVALIFIKAČNÍCH PRACÍ

Kvalifikační práce - dokumenty, které jsou výsledkem studia jednotlivých studijních programů a oborů vysokých škola a vedou k získání vědecké hodnosti či jiné odborné kvalifikace vyššího stupně.

Tato šablona je připravena pro zpracovávání těchto typů vysokoškolských kvalifikačních prací:

habilitační práce
doktorské práce
rigorózní práce
diplomové práce
bakalářské práce

Seznam povinných polí:

Identifikátor (dc:identifier)
Název (dc:title)
Autor (dc:creator)
Akademický titul
Typ studia (thesis:level)
Studijní obor (thesis:discipline)
Instituce přidělující titul (thesis:grantor)
Datum obhajoby (dcterms:dateAccepted)
Datum zveřejnění zdroje (dcterms:dateCreated)
Předmět (dc:subject)

NUŠL typ dokumentu (dc:type)
Jazyk dokumentu (dc:language)
Autorská práva (dc:rights)
Omezení přístupu (nusl:transfer)
Název instituce (nusl:provider)
Popis souboru (nusl:filetype)

4.2.4 Šablona pro zpracování AUTORSKÝCH PRACÍ

Autorské práce - dokumenty, které jsou výsledkem intelektuální činnosti jednotlivých osob.

Tato šablona je připravena pro zpracovávání těchto typů autorských prací:

preprint
referát
monografie

Seznam povinných polí:

Identifikátor (dc:identifier)
Název (dc:title)
Autor (dc:creator)
Datum zveřejnění zdroje (dcterms:dateCreated)
Předmět (dc:subject)
NUŠL typ dokumentu (dc:type)
Jazyk dokumentu (dc:language)
Autorská práva (dc:rights)
Omezení přístupu (nusl:transfer)
Název instituce (nusl:provider)
Popis souboru (nusl:filetype)

4.2.5 Šablona pro zpracování STUDIJNÍCH MATERIÁLŮ

Studijní materiály - dokumenty vznikající při studiu na jednotlivých typech vysokých škol, jejichž autory mohou být pedagogičtí pracovníci nebo studenti.

Tato šablona je připravena pro zpracovávání těchto typů studijních materiálů:

učební osnovy
otázky ke zkouškám
záznamy z přednášek

Seznam povinných polí:

Identifikátor (dc:identifier)
Název (dc:title)
Autor (dc:creator)
Datum zveřejnění zdroje (dcterms:dateCreated)
Předmět (dc:subject)
Instituce přidělující titul (thesis:grantor)
NUŠL typ dokumentu (dc:type)
Jazyk dokumentu (dc:language)

Autorská práva (dc:rights)
Omezení přístupu (nysl:transfer)
Název instituce (nysl:provider)
Popis souboru (nysl:filetype)

4.2.6 Šablona pro zpracování FIREMNÍ LITERATURY

Firemní literatura - dokumenty vznikající v komerčních i nekomerčních organizacích v rámci pracovního procesu a informují veřejnost o produktech, poskytovaných službách a činnostech organizace.

Tato šablona je připravena pro zpracovávání těchto typů firemní literatury:

katalogy produktů
návody

Seznam povinných polí:

Identifikátor (dc:identifier)
Název (dc:title)
Autor (dc:creator)
Datum zveřejnění zdroje (dcterms:dateCreated)
Předmět (dc:subject)
NUŠL typ dokumentu (dc:type)
Jazyk dokumentu (dc:language)
Autorská práva (dc:rights)
Omezení přístupu (nysl:transfer)
Název instituce (nysl:provider)
Popis souboru (nysl:filetype)

4.3 Vkládání dat do digitálního repozitáře

Data se do digitálního repozitáře dostávají nejčastěji přímým vkládáním přes WWW uživatelské rozhraní digitálního repozitáře.

Co je potřeba zajistit, aby bylo možné vkládat data:

- 1) nastavit uživatelské rozhraní,
- 2) připravit pravidla pro fungování repozitáře,
- 3) zaškolit vkladatele.

4.3.1 Nastavení uživatelského rozhraní

Nastavení uživatelského rozhraní závisí na použitém systému a provádí ho správce systému. Obecně lze říci, že je třeba upravit dle potřeb organizace zejména:

Používaný metadatový formát

Předem je třeba určit, jaká data a v jakém formátu se budou ukládat do kterých polí, aby mohla být zajištěna konsistence dat. Tento zdánlivě snadný úkol se během času

může ukázat jako velice svízelný, neboť za běhu repozitáře se můžou vyskytnout nové požadavky nebo se mohou odhalit slabiny původního rozvržení, což může vést k úpravě metadatového formátu. Každá úprava formátu by měla být zpětně reflektována na veškerá stávající data, aby nedocházelo k narušení konsistence dat v repozitáři.

Šablony pro různé typy dokumentů

Různé typy dokumentů mívají různé požadavky na obsah. Jedna universální šablona pro všechny typy dokumentů by mohla být pro zpracovatele záznamu příliš velká, z čehož by vyplývalo nepohodlí při vytváření záznamů a náchylnost k chybám. Je tedy vhodné tyto šablony po uvážení nějak rozumně rozdělit.

Vytvořit sbírky dokumentů

Dokumenty by se měly logicky rozpadat do několika sbírek. Sbírkby by měly tvořit logické celky, které uživatelům umožní se v repozitáři orientovat.

Vytvořit uživatelské skupiny

Je dobré vymezit si určité skupiny uživatelů, které budou mít různá práva. Například hlavní administrátor, správci různých částí repozitáře, běžní uživatelé atd.

Nastavit kontroly dat u formalizovaných polí

Je-li to možné, měla by být nastavena striktní kontrola formátu některých vstupních dat. Typickým příkladem je třeba datum. Pokud se například rozhodneme, že budeme v repozitáři mít data ve formátu ISO 8601 (YYYY, YYYY-MM-DD atd.) měli bychom zamezit tomu, aby se nám do repozitáře dostávala data v jiném formátu.

Doporučení:

V rámci projektu NUŠL vznikla defaultně nakonfigurovaná instalace systému CDS Invenia primárně pro spolupracující organizace, které si ji mohou nainstalovat na vlastním hardwaru a dále upravit dle vlastních požadavků. Tato instalace je volně dostupná. K této instalaci je k dispozici dokumentace, ve které je shrnut jak postup instalace tak i další správa a konfigurace CDS Invenia. Vše je dostupné z webových stránek projektu http://nysl.techlib.cz/index.php/Invenio_lokalni .

4.3.2 Pravidla pro digitální repozitář

Aby mohl digitální repozitář správně fungovat, je potřeba zajistit také legislativní rámec a vymezení obsahu repozitáře, které jsou popsány v první části metodiky, dále definovat metadatový formát a dlouhodobou udržitelnost dat. Na základě těchto podkladů je následně nutné zpracovat následující pravidla:

- nařízení či směrnici upravující obsah digitálního repozitáře, jeho organizační zabezpečení, zpřístupnění digitálních dokumentů a povinností zaměstnanců, viz příklady v kapitole 3.1.6 Vlastní směrnice.
- způsob zápisu jednotlivých prvků v záznamu – je nutný pouze u vlastních prvků metadatového formátu, u převzatých prvků je již způsob zápisu definován (nejlepším řešením je převzít celý metadatový formát)

- seznam podporovaných formátů digitálních dokumentů – systémy většinou nemají nastavená omezení, ale pro dlouhodobou čitelnost digitálních dokumentů je nutné určit omezenou skupinu formátů, které budou podporovány.

Doporučení:

Pro způsob zpřístupnění je efektivní si nechat zpracovat právní expertízu nebo vycházet z již existující viz příklad právní expertíza pro NUŠL níže a proškolit zaměstnance, jak postupovat při uzavírání licenčních smluv s vydavateli v případě preprintů a konferenčních materiálů.

V případě podporovaných formátů je ideální vycházet z formátů, které podporují národní a mezinárodní repozitáře. V případě šedé literatury jsou to formáty preferované v NUŠL, který data dále předává do mezinárodních repozitářů. Aktuální seznam preferovaných formátů je dostupný na stránkách NUŠL <http://nusl.techlib.cz> v části Jak spolupracovat – Informace a návody pro spolupracující instituce.

Příklad:

POLČÁK, Radim. *Digitální zpracování tzv. šedé literatury pro Národní úložiště šedé literatury* [online]. Brno: Právnická fakulta Masarykova univerzita, 2009. Dostupné z WWW: <http://nusl.techlib.cz/images/Pravni_exp_NUSL.pdf>.

FRANTÍKOVÁ, Bohdana. *Pokyny pro zpracování záznamů v systému CDS Invenio* [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nusl.techlib.cz/images/Pokyny_zpracovani_zaznamu.pdf>.

4.3.3 Školení pro vkladatele

Vkladateli jsou většinou pracovníci instituce, kteří prošli školením. V rámci školení je potřeba pracovníky seznámit jednak s používáním uživatelského rozhraní, tedy se systémem digitálního repozitáře, ale také s pravidly pro vyplňování jednotlivých polí. Důležité je vytvořit také manuál pro práci se systémem repozitáře a manuál s pokyny pro zpracování záznamu, ve kterém potřebné informace najdou.

Příklad:

FRANTÍKOVÁ, Bohdana; PEJŠOVÁ, Petra. *Manuál pro vkládání a editaci dat v systému CDS Invenio* [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nusl.techlib.cz/images/Manual_Invenio1.pdf>.

Doporučení:

V šablonách pro vládání dat doporučujeme vytvořit ke každému prvku nápovědu přímo v systému, která bude obsahovat způsob zápisu, příklad a i připomínat všechna pravidla vztahující se k tomuto prvku. Zpracovatelům, to především z počátku, velice usnadní zadávání dat.

5 Dlouhodobé uchování a zpřístupnění šedé literatury

Pro zajištění dlouhodobé ochrany a zpřístupnění dat je potřeba se při budování digitálního repozitáře opřít o standard Reference model for an Open Archival Information System (OAIS). Dokument byl vytvořen v roce 2002 a o rok později byl přijat normou 14721:2003. Norma napomáhá stanovit opatření pro dlouhodobou ochranu digitálních dokumentů (CUBR, 2010). Součástí dokumentu je rámcový model, který nevymezuje softwarové požadavky, ale popisuje jednotlivé prvky a procesy, které se v systému odehrávají. Zároveň je celý systém normou zasazen do širšího kontextu.

V českém jazyce existuje stručný popis obsahu normy v publikaci „Dlouhodobá ochrana digitálních dokumentů“ od Ladislava Cubra viz (CUBR, 2010).

5.1 Doporučené formáty a dlouhodobá udržitelnost digitálních dokumentů

Uchovávání digitálních záznamů je sada procesů a činností zajišťující přístup k informacím, které existují v digitálních formátech. Uchování digitálních dokumentů je v dlouhodobém horizontu probíhající proces. V tomto ohledu jsou informace o uložení digitálního dokumentu důležité pro zajištění jeho životnosti.

Pro dlouhodobou udržitelnost digitálních dokumentů je zapotřebí vytvořit dlouhodobý a bez-chybový postup ukládání digitálních informací, včetně prostředí pro jejich získání a interpretaci. Slovo „dlouhodobý“ zde představuje dostatečně dlouhé období, aby se plán zabýval dopadem měnících se technologií, podpory nových médií a formátů dat, případně měnícími se požadavky uživatelské komunity.

Pro dlouhodobou udržitelnost digitálních dokumentů je potřeba:

- 1) zvolit způsob softwarového zajištění,
- 2) zvolit formáty dokumentů,
- 3) zvolit periodicitu revize formátů v úložišti,
- 4) zapsat údaje do metadatového formátu.

5.1.1 Softwarové zajištění zpřístupnění digitálních dokumentů

Z pohledu rizika zachování digitální informace hrozí v dlouhodobém horizontu zastarání media, případně hardwarového prostředí, ve kterém je médium čitelné. Softwarovému prostředí a formátu dokumentu hrozí ztráta kontextu, v jakém informace vznikla. Zajištění dlouhodobé čitelnosti není řešitelné udržováním zařízení a systémových prostředí z důvodů náročnosti. Je proto třeba zajistit migraci dat. Migrace představuje převedení obsahu do „nového“ formátu – nejlépe otevřeného a standardizovaného. Migrační nástroje zpravidla umožňují převádění dávkově i na vyžádání. Proces migrace může potenciálně přinést rizika nekonzistence informace, ztráty funkcionality či ohrožení kvality, proto je potřeba těmto rizikům předcházet kvalitní přípravou.

5.1.2 Formáty dokumentů

Nejpodstatnějším parametrem pro rozhodnutí je rozšířenost formátu. Bez ohledu na jeho dlouhodobou stabilitu či komerční původ lze konstatovat, že právě pro rozšířené formáty budou dostupné nástroje na případné migrace.

Z toho pohledu mohou být formáty MS Office (DOC, XLS, PPT) sice patřit mezi formáty s nízkou trvanlivostí, avšak budou to formáty snadno migrovatelné. A to ať již do novější verze, tak například do formátu Adobe PDF.

PDF jednoznačně patří mezi obecně doporučené formáty, zejména PDF úrovně A. V případě migrace editovatelných formátů to však může znamenat ztrátu funkcionality.

Microsoft od Office 2007 zavedl Open XML standardy. Formáty zaručují interoperabilitu a transparentnost, které již dnes existují nebo budou vytvořeny. Pro Word, Excel a PowerPoint jsou výchozími příponami souborů v novém formátu DOCx, XLSx a PPTx.

Nové formáty snižují pravděpodobnost poškození souboru a také jeho velikost.

Preferované formáty

Preferované formáty pro šedou literaturu vycházejí jednak z poznatků o obecně rozšířených formátech, ale také z formátů používaných v mezinárodních open access databázích. Preferované formáty je nutné vybírat podle účelu, pro který mají být použity. Z tohoto pohledu je potřeba zohlednit formáty pro zobrazení v prohlížeči, indexaci a archivaci. Konkrétní formáty často splňují více účelů najednou.

Preferované formáty pro zobrazení v prohlížeči

Aby se digitální dokument vždy rychle a bez chyb zobrazil uživateli, je potřeba vybrat vhodné formáty pro zobrazení v prohlížeči. Je tedy nutné vybrat formáty, které jsou rozšířené, bez nutnosti instalace dalších komponent na uživatelský počítač. V případě obrazu, zvuku a videí je navíc důležité vybrat formáty, které lze komprimovat v požadované kvalitě, aby velikost digitálního dokumentu byla kapacitně možná pro přenos dat.

Preferované formáty pro indexaci

Indexace prakticky znamená extrahování čistého textu z dokumentu a zaznamenání výskytu řetězců písmen do indexu. Z pohledu udržitelnosti informace a čitelnosti souboru lze formáty považovat za vhodné pro indexaci, lze-li z nich získat textovou informaci.

Preferované formáty pro záložní archivaci

V případě potřeby kopie pro záložní archivaci je nejvhodnější formát PDF/A-1a. Část 1 PDF/A ISO standardu [ISO 19005-1:2005] je formou Adobe PDF verze 1.4. Formát byl vytvořen se záměrem dlouhodobého uchování dokumentů, pro které je používán formát PDF. Úroveň A (PDF/A-1a) indikuje kompletní shodu s požadavky ISO 19005-1.

V následující tabulce jsou vypsány preferované formáty pro spolupráci s NUŠL a další doporučené formáty. Zde je ještě důležité upozornit, že tyto doporučené formáty jsou aktuální v roce vzniku metodiky (2011). V následujících letech je nutné vždy ověřit aktuální stav na stránkách NUŠL <http://nusl.techlib.cz/>.

Typ dokumentu	Preferované formáty	Doporučené formáty
Textové dokumenty	PDF	TXT, DOC-DOCx, XML (HTML)
Tabulkové dokumenty	PDF	CSV, XLS-XLSx
Prezentace	PDF	(PPT -PPTx)
Grafické formáty	JPG	JP2, PNG, TIFF
Zvuk	MP3	AVI
Video	MP4	MPEG-1 (-2), AVI

Po té, co si instituce určí preferované formáty pro uchovávání dokumentů, lze digitální dokumenty v těchto formátech ukládat přímo do repozitáře. V případě jiných formátů je vhodná migrace na preferovaný formát ihned při vkládání do repozitáře nebo revize úložiště v pravidelných intervalech.

5.1.3 Revize formátů v úložišti

Následně je nutné si určit periodicitu pravidelných cyklů kontroly formátů v úložišti a migrace na preferovaný formát. Zároveň je potřeba si předem stanovit v jakých cyklech bude revidována platnost preferovaných formátů či si určit důvody, které budou mít vliv na změnu preferovaných formátů. Důvody mohou být různého charakteru např. vydání nové verze software, nového interního předpisu (v případě sjednocování formátů v rámci organizace), změna uživatelských preferencí nebo preferovaných formátů v národní či mezinárodní databázi, kam jsou data zasílána. Pokud nenastane žádný z vnějších vlivů na změnu preferovaných formátů lze doporučit revizi každých 12 měsíců. Instituce spolupracující s NUŠL budou o změnách preferovaných formátů informovány.

5.1.4 Zápis informací o formátu v metadatovém záznamu

Pro zápis informace o formátu digitálního dokumentu v metadatovém záznamu je vhodné použít atribut u prvku (elementu) obsahujícího URL digitálního dokumentu, aby bylo možné zapsat a identifikovat formáty u jednotlivých digitálních dokumentů připojených k metadatovému záznamu. V případě spolupráce s NUŠL použijte atribut `nusl:fileType`.

Jako příklad může posloužit prvek ve formátu NUŠL:

`nusl:fileType`

Nově vytvořený atribut pro prvek (element) `nusl:transfer`.

Popis digitálního dokumentu. Specifikuje formát digitálního dokumentu, který je přiložen k metadatovému záznamu. Atribut obsahuje výběr z těchto hodnot:

- Plný text
- Prezentace
- Obrazová část
- Videozáznam
- Audiozáznam
- Posudek vedoucího práce
- Oponentura
- Příloha
- Obrázek
- Tabulky

Na základě obsahu pole `nusl:fileType` systém CDS Invenio vytváří pole `dcterms:medium`, pole je využíváno pro vytvoření navigace **Formát** v uživatelském rozhraní www.nusl.cz. Pole nabývá těchto hodnot:

- text/pdf
- image/jpg
- video/mp4
- audio/mp3

Zobrazení NUŠL formát 1.0

```
<nusl:transfer nusl:fileType="Plný text">http://invenio.nusl.cz/record/42899/files/idr-69_1.pdf</nusl:transfer>
<nusl:transfer
nusl:fileType="Prezentace">http://invenio.nusl.cz/record/42899/files/idr-69_2.pdf</nusl:transfer>
<nusl:transfer
nusl:fileType="Prezentace">http://invenio.nusl.cz/record/42899/files/idr-69_3.pdf</nusl:transfer>
<nusl:transfer
nusl:fileType="Videozáznam">http://invenio.nusl.cz/record/42899/files/idr-69_4.mp4</nusl:transfer>
```

Doporučení:

Při exportu záznamů pro NUŠL je nutné používat preferované formáty podle tabulky výše. Aktuální stav formátů bude vždy k dispozici na webových stránkách projektu NUŠL <http://nusl.techlib.cz/>. U preferovaných formátů lze doporučit revizi každých 12 měsíců, pokud nenastanou důvody ke změně zmíněné výše. Instituce spolupracující s NUŠL budou o změnách preferovaných formátů automaticky informovány. Důležité je neopomenout zapsat informaci o formátu v metadatovém záznamu.

5.2 Persistentní identifikátor

Persistentní identifikátor (dále jen PID) je jednoznačný a trvalý identifikátor zdroje, který se v průběhu času nemění. Identifikátor jednoznačně identifikuje zdroj v rámci repozitáře zatímco PID identifikuje zdroj globálně. V současné době existuje ve světě několik systémů PID. Mezi nejznámější identifikátory fyzických dokumentů patří ISBN pro knihy a ISSN pro časopisy. V případě ISBN systém funguje tak, že každé vydání a obměna (kromě přetisku) knihy dostane jedinečné číslo ISBN. V digitálním věku stále roste potřeba takového jedinečného identifikátoru také pro digitální dokumenty. Dnes již běžnými identifikátory používanými ve vědecké komunikaci jsou DOI (používá se především pro časopisecké články v digitální podobě), HANDLE, URN:NBN a další.

Tyto identifikátory bývají často strukturované (např. URN:NBN:fi-fe19981001) a tím jsou částečně čitelné i pro člověka, většinou se však používá kódové označení (např. využití systémového čísla jako identifikátoru). Aby nedocházelo ke kolizím mezi jednotlivými lokálními repozitáři, je nutné, aby tyto identifikátory byly světově unikátní, nejlépe přidělovány nějakou nadřízenou autoritou.

Pro zavedení PID je potřeba:

- 1) stanovit kritéria,
- 2) vybrat vhodný PID,
- 3) zajistit jednoznačnost,
- 4) zajistit perzistentnost.

5.2.1 Stanovení kritérií

Instituce, která se rozhodne používat systém PID, musí pro něj vytvořit organizační podmínky a na praktické úrovni si zodpovědět řadu otázek. Hlavním úkolem je navrhnout samotný identifikátor. Jednou z hlavních otázek je, zda se organizace připojí k existující implementaci systému PID nebo zda bude implementovat vlastní systém včetně resolveru (systému pro přiřazení konkrétního záznamu/digitálního dokumentu k identifikátoru, realizovanému např. zobrazením odpovídajícího dokumentu/ digitálního dokumentu z repozitáře). K návrhu identifikátoru vede cesta přes zodpovězení následujících otázek.

Typ identifikátoru

Perzistentní identifikátor může představovat buď náhodný výběr znaků bez jakýchkoliv

souvislostí, které neobsahují žádné informace o objektu a určuje je náhodný řetězec znaků (anglicky „dumb number“). Nebo se může jednat o předem daný systém, který využívá identifikace např. země, repozitáře, kolekce a dokumentu. Takovýto systém inteligentních identifikátorů může mít různou mírou složitosti a konkrétnosti. Obecně lze říci, že lidem se lépe pamatují a používají inteligentní identifikátory s vestavěnou mnemotechnikou, než bezvýznamná posloupnost znaků, ale pro strojní zpracování je toto hledisko irelevantní. U inteligentních identifikátorů lze také snáze dosáhnout jedinečnosti, a to nejen mezi různými organizačními jednotkami, ale zejména globálně. Generování náhodných řetězců znaků musí být naopak pečlivě kontrolováno a sledováno v celé organizaci s cílem zajistit jedinečnost, kterou však již nelze zajistit i mimo organizaci. Inteligentní identifikátor může obsahovat relační informace, kdy menší komponenty jsou určeny odkazem na větší subjekty, nebo sbírky, kterých jsou součástí.

Varianty identifikátorů:

- Náhodný řetězec znaků („dumb number“)
- Inteligentní identifikátor

Hierarchie

Jedním z nejběžnějších použití inteligentních identifikátorů v knihovním světě je začlenění

relačních informací, které zobrazují organizaci a hierarchie digitálních sbírek. U menších složek, jako jsou digitální obrazy stránek nižší úrovně, je součástí PID identifikace nadřazené složky. Tato forma identifikace se hodí zejména pro digitální reprezentace fyzických sbírek, i když se může také použít k zobrazení hierarchie digitální sbírky, jako jsou čísla časopisů a jejich články.

Tento systém hierarchie ale není nutný v případě, že každá úroveň např. časopisu (titul,

ročník, číslo, stránka) má vlastní identifikátor. V těchto případech není třeba, aby syntaxe identifikátoru poskytovala na první pohled informaci o hierarchii. Rizikem při použití hierarchie pro tvorbu PID je situace, kdy se např. změní strukturování dokumentů a hrozí tak změna identifikátoru, ke které v praxi nesmí dojít (podmínka perzistence).

Varianty:

- Podporovat hierarchii při tvorbě PID
- Nepodporovat hierarchii při tvorbě PID

Granularita (úroveň odkazu)

Rozhodnutí o úrovni detailu odkazu, na které budou PID přidělovány, závisí na vnímání potřeb ukládaného materiálu. Pro mnohé potřeby vystačí PID na nejvyšší úroveň granularity, který může být reprezentován např. webovou stránkou, která slouží jako vstupní bod např. na článek v časopise, či odkaz na konkrétní kapitoly či v knize. Nicméně některá použití mohou vyžadovat jemnější úroveň detailů. Granularitu si musí určit každá instituce pro jednotlivé typy dokumentů sama.

Příklad granularity:

- Celý dokument
- Kapitoly
- Stránky
- Obrázky

Verze

Verze jednoho dokumentu se mohou lišit různými způsoby, může mít jiný obsah, jiný formát, nebo jiné rozlišení ve stejném formátu např. u videí s různým datovým tokem. Každá verze objektu, pokud sledujeme verze, vyžaduje samostatný PID. Vztah mezi verzemi může být vyjádřen v identifikátoru pomocí kódu verze nebo datováním verze nebo kódem typu verze, nebo v metadatech. Je tedy nutné zvážit, zda a jak zaznamenávat vztahy mezi jednotlivými verzemi.

Varianty verzování:

- Nesledovat verze
- Verze pro odlišný obsah
- Verze pro odlišný formát
- Rozlišení ve stejném formátu

Zajišťující autorita

Na zajišťující autoritě závisí nejdůležitější úkol pro úspěšný systém PID, a to organizační zajištění. Zajišťující autorita garantuje perzistentnost identifikátoru, čímž na sebe bere zodpovědnost za jeho dlouhodobé přetrvání. Zajišťující autorita zajistí správu resolveru pro PID (údržbu registru vztahů mezi PID a reprezentovanými digitálními objekty, aktuálnost tohoto registru) a přidělování jmenných prostorů jednotlivým institucím. Přidělování samotných PID může zajišťovat centrální instituce nebo instituce s přiděleným jmenným prostorem sama (viz dále). Zajišťující autorita nemusí být specializovaná na šedou literaturu, ale musí být schopna zajistit vybrané schéma PID a jeho perzistentnost.

Varianty zajišťující autority:

- Vlastní instituce (např. PURL - perzistentní URL adresa)

- Správce zavedeného systému (např. Handle, URN:NBN a další)

Finanční hledisko

Pro management digitálního repozitáře bude jistě finanční hledisko velmi významné, proto je důležité ho neopomenout. Např. některé identifikátory je možné využívat bez zvýšených finančních nákladů (např. URI), jiné jsou zpoplatněny (např. systém Handle).

Variety nákladů na PID:

- Pořizovací náklady, dále zdarma
- Pořizovací a udržovací náklady
- Pravidelné roční příspěvky
- Zdarma

viz <http://nysl.techlib.cz/index.php/Identifikatory>

5.2.2 Výběr vhodného PID

Poté, co si instituce odpoví na otázky výše, je na řadě porovnat definované požadavky s možnostmi, které se nabízejí, a to buď vybrat z již existujících systémů PID, nebo implementovat vlastní řešení. Obecně lze říci, že v digitálních repozitářích šedé literatury není potřeba klást speciální nároky na hierarchii, granularitu a verzování. Naopak, je důležité zvolit důvěryhodnou zajišťující autoritu, zvážit finanční hledisko a zajistit jednoznačnost a perzistentnost identifikátoru. Perzistentní identifikátory jsou potřebné k vytvoření stabilní a spolehlivé infrastruktury. To se netýká technických podrobností, ale hlavně dohod na organizační úrovni. Z důvodu ověřeného organizačního zajištění je tedy nejlepším řešením vybrat některý z PID jako jsou URN:NBN, Handle a DOI provozované registračními autoritami.

V rozhodování může pomoci také doporučení a požadavky mezinárodních repozitářů, ke kterým je vhodné se připojit. V případě šedé literatury je vhodné respektovat doporučení DRIVER Guidelines, které uvádějí jako optimální postup identifikovat zdroj pomocí řetězce čísel odpovídajících formálnímu identifikačnímu systému. „Příklady formálních identifikačních systémů zahrnují Uniform Resource Identifier (URI) (včetně Uniform Resource Locator (URL), Digital Object Identifier (DOI) a URN:NBN). Ideálním využitím tohoto prvku je použití přímého odkazu nebo odkazu na výchozí stránku (trvalá URL) z dc:identifier v metadatovém záznamu na digitální zdroj nebo výchozí stránku. Trvalý odkaz může být na záznam nebo na soubor. Může to také být URL nebo přesměrovaná URL jako PURL, HANDLE nebo jiné mechanismy mezinárodního řešení.“¹²

Dalším je doporučení OpenAIRE Guidelines, které definují identifikátor jako „URI schéma, které je nalinkováno na persistentní identifikátor (URN, handle, DOI), plný text nebo výchozí stránku. Identifikátor by měl být jedinečný.“¹³

Doporučené existující systémy PID

Archival Resources Key (ARK)

ARK je URL vytvořené tak, aby zajistilo perzistentní a dlouhodobý přístup k informačním objektům. Pomocí ARK je možné identifikovat objekty všeho druhu:

¹² DRIVER Guidelines [online], s. 52. [cit. 2011-10-10].

¹³ OpenAIRE Guidelines 1.1 [online]. 2010, s. 10. [cit. 2011-10-10]

digitální dokumenty, databáze, obrázky, software a webové stránky, stejně jako fyzické předměty (knihy, sochy apod.) a též nehmotné objekty (chemikálie, nemoci, představení apod.).

ARK podporuje trvalou identifikaci, která je nezbytná a užitečná, protože jak protokoly (např. http a ftp) používané pro přístup k objektům, tak stránky, které objekty hostí, se mohou měnit. Záměrem je vytvořit stabilní „jméno“ nebo odkaz, který bude trvale spojen s konkrétním objektem.

Další informace naleznete na <https://wiki.ucop.edu/display/Curation/ARK>

Digital Object Identifier (DOI®)

Identifikátor digitálního objektu (DOI) je systém pro identifikaci obsahu objektů v digitálním prostředí. DOI jméno jsou přiřazovány entitám pro použití v digitálních sítích, které slouží k poskytování aktuálních informací včetně toho, kde jsou k dohledání na internetu. Informace o digitálním objektu se může v čase měnit, včetně toho, kde je k dohledání, ale jeho DOI jméno se nikdy nezmění. Systém DOI poskytuje rámec pro trvalou identifikaci, řízení intelektuálního obsahu, správu metadat atd. Systém DOI splňuje mezinárodní normu ISO. DOI systém je řízen Mezinárodní DOI nadací, sdružení nabízí otevřené členství pro komerční i nekomerční partnery. Prostřednictvím registračního systému DOI agentury již bylo přiřazeno více než 50 milionů DOI jmen.

<http://www.doi.org/>

Uniform Resource Name: National Bibliography Number (URN:NBN)

Každá národní knihovna využívající tento identifikátor používá svůj vlastní NBN (National Bibliography Numer) řetězec nezávisle na jiných národních knihovnách, s tím, že neexistuje globální orgán, který toto kontroluje. Z tohoto důvodu jsou NBN jedinečné pouze na národní úrovni. Při použití jako URN, musí být doplněny NBN řetězce kontrolou prefixu, jako je kód země. Tyto předpony zaručují jedinečnost NBN, založeném na URN, v celosvětovém měřítku.

<http://wiki.surffoundation.nl/display/standards/URN-NBN>

Český URN:NBN resolver

<http://resolver.nkp.cz/urnnbn/>

Handle

„Systém Handle směřuje identifikátory nezávisle na DNS a nemusí spoléhat na URL. Může je ale využívat prostřednictvím proxy serverů, které identifikátor na URL přeměrují, a stále zajišťují a aktualizují vazby mezi nimi. Systém je spravován globálním serverem, který spravuje databázi institucí a korporací, které identifikátory HDL přidělují. Založila ho americká Korporace pro národní výzkumné iniciativy CNRI (Corporation for National Research Initiatives). Původně byl navržen pro přístup k technickým zprávám uloženým v digitálním archivu. Díky systému Handle je digitálnímu objektu umožněno, aby byl identifikován, byl přístupný a chráněný. Digitální objekt i v tomto případě může být spojen s metadaty.“(Šímová, 2010)

<http://www.handle.net/>

Persistent Uniform Resource Locators (PURL)

Jsou webové adresy, které se chovají jako trvalé identifikátory v prostředí dynamicky se měnící webové infrastruktury. Místo odkazování přímo na webové zdroje, PURL poskytuje úroveň dereference, které umožňují změny základních webových adres zdrojů v čase, aniž by to mělo negativní dopad na nich závislých systémů. Tato funkce zajišťuje kontinuitu odkazů pro síťové prostředky.

<http://purl.oclc.org/docs/index.html>

Uniform Resource Identifier (URI)

Jednotný identifikátor zdroje URI patří mezi jedny z nejstarších identifikátorů pro digitální objekty. Jeho popis je uveden v normě RFC 3986. URI je definován jako řetězec znaků, který slouží k identifikaci (nebo lokalizaci, vyhledávání, směřování, odkazování apod.) Použit může být pro jakýkoli objekt. Webovou stránku, databázi, službu, dokument, fyzický objekt, metadatový záznam. URI je obecným rámcem identifikátorů informačních zdrojů. Skládá se ze dvou podtypů URN a URL. (Šímová, 2010)

<ftp://ftp.rfc-editor.org/in-notes/rfc3986.txt>

5.2.3 Zajištění jednoznačnosti

Nejdůležitější požadavek je, aby se jeden identifikátor vztahoval vždy jen k jednomu záznamu. Žádoucí je samozřejmě také, aby jeden záznam měl pouze jeden identifikátor.

Identifikátory musí být unikátní v rámci digitálních zdrojů organizace a v kombinaci s URL unikátní i celosvětově. V případě vlastního řešení vytvoříme identifikátor nejjednodušeji z kombinace systémového čísla, které repozitář každému záznamu přiděluje a URL adresy organizace. Alternativně lze uvažovat i o jiné kombinaci atributů záznamu (například rok, autorita a nějaké kódové označení), ovšem tyto postupy bývají komplexnější a nemusí být vždy zaručena jednoznačnost.

5.2.4 Zajištění perzistence

Je důležité, aby zdroj identifikovaný PID nebyl nikdy přesunut nebo odstraněn, aniž by došlo k aktualizaci informací o jeho umístění v registru perzistentních identifikátorů. Instituce musí udržovat propojení aktuálního umístění zdroje na PID.

Cílem identifikátoru je, abychom se pomocí něho dostali k odpovídajícímu záznamu. V případě, že jako součást identifikátoru použijeme systémové číslo, tak toto číslo bývá často i součástí URL konkrétního záznamu. Z důvodu perzistence však není vhodné publikovat URL, která přímo odkazují na záznam. Došlo-li by totiž k přesunutí repozitáře pod jinou doménu, změnila by se i všechna URL k záznamům. Je vhodné použít zde nějakého prostředníka (např. webovou službu, resolver), kterému zadáme PID a on nás odkáže na aktuální umístění odpovídajícího záznamu a publikovat URL, která odkazují na tento resolver.

Pro vytváření PID a směřování k záznamům je možné využít již existující služby (Handle, DOI aj.). Tyto nám poskytnou jednotnou tvorbu PID a způsob, jakým se dostaneme k záznamu. Tyto služby se liší především způsobem přesměrování (Handle využívá persistentní URL, DOI má interaktivní rozhraní) a také cenou za

služby. Avšak i při využití těchto služeb je důležité, aby nedocházelo ke změnám identifikačních údajů v repozitáři (např. k přečíslování záznamů).

Příklad:

V projektu NUŠL se identifikátor vztahuje k metadatovému záznamu, nikoliv k vlastním plným textům. Odkazy na plné texty jsou součástí metadatového záznamu. Struktura identifikátoru je nusl-[číslo] (např. nusl-28). Číslo je generováno repozitářem, do kterého se záznamy ukládají. Povinností administrátora repozitáře je zaručit, že v průběhu času nedostanou dva různé záznamy stejné číslo (to se může stát například při migraci dat).

Globální identifikaci lze jednoduše řešit například pomocí URL adresy, jejíž jednou součástí je identifikátor.

Příklad:

<http://adresa.repozitare.cz/zaznam/nusl-28>

Nevýhoda této jednoduché myšlenky je v tom, že adresa repozitáře se může změnit a tím by se změnilo i URL záznamu. V případě takovéto změny je zapotřebí zajistit odpovídající přesměrování na novou adresu z původní domény. Pro projekt NUŠL byla zřízena webová služba, která persistentnost URL odkazu zajišťuje. Adresa služby je

Syntaxe:

<http://www.nusl.cz/ntk/> + identifikátor

Příklad URL:

<http://www.nusl.cz/ntk/nusl-28>

Tento odkaz představuje trvalý odkaz na záznam. Pokud by se adresa repozitáře změnila, stačí jen přenastavit parametry této služby a URL zůstane stejné, stále platné.

NUŠL se v současné době snaží o přijetí identifikátoru URN:NBN, jehož přiděluje autoritou je Národní knihovna ČR. Při jeho zavádění se počítá i s využitím zmiňované webové služby, která uživatele odkáže na požadovaný záznam.

5.3 Zpřístupňování a předávání dat

Hlavním cílem každého digitálního repozitáře je zpřístupnění jeho obsahu, kterému napomáhá předávání dat o digitálních dokumentech do obsahově či jinak zaměřených repozitářů na národní či mezinárodní úrovni. V této části metodiky je popsán způsob, jak přehledně zobrazovat data uživatelům a návod, jak je předávat pomocí protokolu OAI-PMH.

5.3.1 Zpřístupňování dat v digitálním repozitáři

Záznamy by měly být rozděleny v digitálním repozitáři do sbírek podle nějakého logického kritéria tak, aby to uživateli umožňovalo jednoduchou orientaci.

Pro stavbu sbírek v systému je potřeba využít nebo specifikovat odpovídající prvky metadatového formátu. Sbírkou lze stavět podle různých hledisek a tato hlediska je užitečné vzájemně kombinovat. Nejčastější je využíváno stavění podle typu

dokumentu, dále je možné stavět sbírky podle věcného hlediska, jednotlivých projektů a činností dané instituce, nebo podle organizační struktury, resp. jednotlivých oddělení popř. odborů konkrétní organizace.

Sbírky dle typu dokumentu

Pro stavbu této sbírky je vhodný prvek dc:type.

Využití tento způsob řazení dokumentu je možné doporučit v případě, že je předem známý konečný počet a rozsah dokumentů určených ke zpracování v repozitáři. Je pečlivě zpracována typologie dokumentu a množství typů dokumentů je přiměřené z hlediska přehlednosti v systému.

Rizikem při využití této typologie jako hlediska řazení dokumentu do sbírek je, že pokud by typologie byla připravena příliš úzce, některé typy dokumentu by byly zastoupeny pouze několika dokumenty, resp. záznamy o dokumentech. Tato skutečnost pak nutně vede k tomu, že vzniká potřeba vytvořit obecnou sbírku, tzv. ostatní, do které jsou pak řazeny dokumenty, které se do stanovené typologie „nevešly“, což následně vede ke snížení přehlednosti takto tvořených sbírek.

Sbírky dle projektů a činností

Pro toto řazení je potřeba specifikovat vlastní nový prvek a doplnit ho do zvoleného formátu.

Obsahem prvku budou názvy jednotlivých projektů, na kterých instituce pracuje a kontinuálně vykonávané významné činnosti, např. rekvalifikační kurz, digitalizace apod.

Toto řazení je vhodné, především pokud je práce na projektech pro instituci hlavní náplní a jsou přesně dány další činnosti organizace. Za riziko tohoto řazení lze považovat, že výstupů z projektů a činností nebude mnoho. Toto řazení je možné zkombinovat s některou z dalších zmiňovaných možností.

Sbírky dle organizační struktury

Je potřeba specifikovat vlastní nový prvek, který bude součástí metadatového formátu a ve kterém bude uložena informace o instituci, odboru, nebo oddělení, ve kterém dokument vznikl. Možné využít, pokud je organizační struktura dobře stanovena a nepředpokládají se v budoucnosti významné změny v hierarchii organizace. Rizikem využití tohoto řazení je skutečnost, že dokument mohl vzniknout ve více odděleních – původce pak není jasně známý, a dále riziko, že především v oblasti státní správy jsou reorganizace poměrně častým jevem.

Sbírky dle věcného hlediska

Pro stavbu této sbírky je vhodný prvek dc:subject.

Stavění sbírky z věcného hlediska je možné doporučit v případě, kdy je k věcnému popisu využíván některý řízený slovník (heslář, tezaurus), nebo klasifikační slovník. Nelze využít, pokud je věcný popis tvořen pouze využitím volně tvořených klíčových slov. Rizikem tohoto způsobu stavění sbírek je volba úrovně (hloubky) popisu tak, aby se neztratila přehlednost struktury.

Příklad:

V NUŠLu jsou tato kritéria dvě – typ dokumentu a instituce, ze které tento dokument pochází. Obě rozdělení jsou navíc uspořádána hierarchicky pro lepší přehlednost.

Doporučení:

Možnosti vyhledávání je třeba nastavit tak, aby odpovídaly datům. Je třeba zamyslet se nad tím, jaká pole budou nejčastěji prohledávána uživateli a podle toho nastavit volby vyhledávání a indexy.

Zobrazení záznamů by vždy mělo odpovídat dané situaci. Pokud zobrazujeme několik záznamů najednou (například při vyhledávání), je dobré zobrazit jen nejdůležitější informace. Při detailním zobrazení konkrétního záznamu není nutné zobrazovat naprosto veškeré informace, ale mělo by zde být vše podstatné v přehledné formě.

Zpřístupňování digitálních dokumentů

Zpřístupňování obsahu repozitáře není pouze zobrazením metadat o digitálních dokumentech nebo klasických dokumentech šedé literatury, ale také zprostředkování obsahu těchto dokumentů.

V případě, že metadatový záznam doprovází digitální dokument je bezpodmínečně nutné, aby odkaz na tento dokument byl trvalý a předešlo se tak problémům při aktualizaci jednotlivých linků.

Pokud digitální dokument není k dispozici, musí být součástí záznamu informace o dostupnosti, tj. v jaké instituci a pro jakou skupinu uživatelů je dokument dostupný, dále informace o službě prostřednictvím, které je možné dokument získat a kontaktní informace na instituci (osobu), která službu poskytuje (tj. email, adresa apod.).

Doporučení:

Vzhledem k tomu, že lze předpokládat častější změny v kontaktních informacích, doporučujeme pro jejich zaznamenávání vytvořit samostatnou databázi a údaje nezaznamenávat v metadatech.

5.3.2 Předávání dat

Předávání dat je důležité pro spolupráci s dalšími digitálními repozitáři na národní i mezinárodní úrovni. V dnešní době již většina repozitářů vyžaduje předávání dat pomocí OAI-PMH¹⁴ protokolu za účelem periodického, automatického a konzistentního sklizení dat. Pokud není možné použít protokol OAI-PMH, pak vždy lze data nějakým způsobem zasílat dávkově (např. e-mailem, přes FTP apod.). Tento postup však vyžaduje větší množství "manuální práce" a pokud to lze, je lepší se mu vyhnout. Většina národních a mezinárodních repozitářů možnost dávkového předávání dat neposkytuje.

OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) je internetový protokol vyvinutý Open Archives Initiative. Používá se pro harvestování (sklizení) metadatových záznamů z určitého repozitáře. Povinný formát pro harvestování je Dublin Core, volitelně je možné používat i další formáty. OAI-PMH přenáší XML soubory protokolem HTTP.

¹⁴ <http://www.openarchives.org/OAI/openarchivesprotocol.html> [cit. 2010-10-10]

Pro implementaci OAI-PMH v digitálním repozitáři je potřeba zajistit:

- 1) povinné a volitelné náležitosti OAI-PMH serveru,
- 2) nakládání se smazanými záznamy,
- 3) povinné prvky pro šedou literaturu,
- 4) v případě vysokých škol povinné prvky pro VŠKP.

Povinné a volitelné náležitosti OAI-PMH serveru

OAI-PMH server je provozován na specifické base URL adrese.

<http://adresa.repozitare.cz/oai-script>

Serveru na této adrese se posílají protokolem HTTP proměnné, které mohou být typu POST (v hlavičce HTTP požadavku) nebo GET (v URL adrese). V praxi se většinou používá GET. Kombinace proměnných a jejich hodnot pak vyvolávají specifické odezvy.

Proměnné mohou být následující:

- verb - určuje, jaká akce se bude provádět
- metadataPrefix - určuje, v jakém formátu budeme data harvestovat
- resumptionToken - pro navázání přerušeno harvestování či pro získání následující dávky záznamů
- from - záznamy od jakého data chceme harvestovat
- until - záznamy do jakého data chceme harvestovat
- identifier - OAI identifikátor záznamu
- set - určitá skupina záznamů. Skupin může být více, nemusí být disjunktní a mohou být i hierarchické

Společná struktura všech odpovědí vypadá takto:

```
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
<responseDate>YYYY-MM-DDTHH:mm:ssZ</responseDate>
<request verb="[hodnota verb]">http://adresa.repozitare.cz/oai</request>
[Vlastní obsah]
</OAI-PMH>
```

Vlastní obsah se potom liší dle hodnoty proměnné verb a další proměnných. Vždy je však uzavřen v elementu, který odpovídá hodnotě proměnné verb (např. <Identify>, <ListMetadataFormats>, popř. <error>)

Proměnná verb

Proměnná verb je vždy povinná, její absence by měla vracet příslušnou chybu. Proměnná verb může nabývat následujících hodnot (záleží na velikosti písmen):

- Identify - identifikace repozitáře, obsahuje základní informace o repozitáři, jako URL adresu, verzi OAI-PMH rozhraní, obsah repozitáře, e-mail administrátora repozitáře a mnohé další

- ListMetadataFormats - vrátí výpis všech možných formátů, ve kterých lze metadata harvestovat
- ListSets - vrátí výpis všech dostupných metadatových setů
- GetRecord - Vrátí jeden konkrétní záznam. Pro verb=GetRecord jsou povinné také proměnné identifier obsahující OAI identifikátor záznamu a metadataPrefix
- ListRecords - vrátí dávku záznamů. Pro verb=ListRecords jsou povinná proměnná metadataPrefix a volitelné proměnné set, from a until. Struktura záznamu vypadá takto:

```
<record>
  <header>
    <identifier>oai:adresa.repozitare.cz:SYSNO</identifier>
    <timestamp>YYYY-MM-DDTHH:mm:ssZ</timestamp>
    <setSpec>set</setSpec>
  </header>
  <metadata>
    ...
  </metadata>
</record>
```

- V identifikátoru v části SYSNO by se měl nacházet unikátní identifikátor v rámci repozitáře, většinou číslo, které repozitář záznamu přiděluje. Struktura identifikátoru musí odpovídat syntaxi URI. Element timestamp obsahuje datum a čas poslední modifikace záznamu v UTC formátu a řídí se podle něj časové omezení definované proměnnými from a until. Element setSpec obsahuje název setu, ve které se záznam nachází.
- ListIdentifiers - vrátí pouze hlavičky záznamů. Povinná je zde proměnná metadataPrefix. Volitelné proměnné jsou from, until a set.

Proměnná metadataPrefix

Určuje metadatový formát záznamů. Každý repozitář by měl umožňovat harvestovat záznamy ve formátu DC (metadataPrefix=oai_dc), ale počet a druh dalších formátů není nijak omezen.

Proměnné from a until

Proměnné from a until určují časové rozmezí (od-do), z jakého budeme data harvestovat. Řídícím datem je zde datum poslední modifikace záznamu. Toho se využívá například při průběžném harvestování repozitáře, kdy stačí, aby si harvestující repozitář uchoval datum posledního harvestování a následující stažení dat provede s omezení from=(datum posledního harvestování). Všechna data jsou ve formátu UTC, tedy YYYY-MM-DDTHH:mm:ssZ, např. 2011-01-22T18:52:31Z. Není nutné zadávat přesný čas, lze zadat jen rok, měsíc a den.

Proměnná resumptionToken

Záznamů v repozitáři může být velké množství (statisíce, miliony) a není tedy vhodné harvestovat všechny najednou. Proto záznamy mohou být rozděleny do dávek. Doporučená velikost dávky je 100-500 záznamů, někdy jsou použity dávky o velikosti až 1000 záznamů. Nevejdou-li se všechny záznamy do jediné odpovědi, tak je za poslední záznam v dávce zařazen tzv. Resumption Token. Jedná se o unikátní sérii

znaků (řádově desítky), pomocí kterých identifikujeme navazující dávku. Resumption Token má omezenou platnost (obvykle cca 2 dny). Resumption Token můžeme obdržet pouze tehdy, pokud proměnná verb nabývá hodnot ListRecords nebo ListIdentifiers. Následující dávku si vyžádáme dalším HTTP požadavkem s příslušnou hodnotou proměnné verb a resumptionToken, ale už nezadááme proměnné metadataPrefix ani set (ty jsou dané Resumption Tokenem).

```
...
</metadata>
</record>
<resumptionToken expirationDate="YYYY-MM-DDTHH:mm:ssZ">ff4990cb9804df0721439fbd73cbfefaf</resumptionToken>
</ListRecords>
</OAI-PMH>
```

Nakládání se smazanými záznamy

Se smazanými záznamy může být nakládáno třemi způsoby:

- **no** - záznam ze setu zmizí
- **persistent** - v setu zůstane pouze hlavička záznamu, element header bude mít atribut status="deleted" a celý element <metadata> chybí. Pokud je použita tato metoda nakládání se smazanými záznamy, musí repozitář zaručit, že každý smazaný záznam bude v setu trvale vystaven tímto způsobem.
- **transient** - stejné jako persistent s tím rozdílem, že se repozitář nezavazuje k tomu, že smazané záznamy budou vystaveny trvale či že budou vystaveny všechny.

Povinné prvky pro šedou literaturu

Povinné prvky NUŠL v OAI-PMH

Použití jednotlivých prvků Dublin Core v XML záznamech šedé literatury zprostředkovaných prostřednictvím OAI-PMH serveru se řídí standardy Dublin Core (formální zápis prvků) a NUŠL (obsahová náplň prvků). Tato metodika stanovuje pro potřeby sdílení záznamů šedé literatury tyto povinné prvky standardu NUŠL:

Prvky povinné pro všechny typy dokumentu:

- dc:identifier – jednoznačný trvalý identifikátor dokumentu poskytovatele metadat. Hodnota povinného atributu typu identifikátoru zahrnuje libovolné schválené schéma podle seznamu permanentní identifikátorů IANA. Při zápisu identifikátorů se postupuje podle doporučení DCMI pro používání identifikátorů zdroje v metadatech DC.
- dc:title – název dokumentu
- dc:title (nusl:typeTranslated) – název dokumentu
- dcterms:dateCreated – datum zveřejnění zdroje
- dc:subject – věcný popis dokumentu
- dc:creator – zápis autora – fyzické osoby, která je primárně odpovědná za vytvoření dokumentu, a to ve formátu Příjmení, Jméno.
- dc:type – typ zdroje (obsahové hledisko, forma) pro standardu NUŠL
- dc:language – jazyk zdroje
- dc:rights – textová informace o právech k využívání zdroje.

Prvky povinné pro Konferenční materiály:

- nusl:event – název konference, akce
- nusl:eventPlace – místo konání konference, akce
- nusl:eventDate – datum konání konference, akce

Prvky povinné pro VŠKP:

- thesis:name – zkratka jména akademického titulu nebo vědecko-pedagogické hodnosti
- thesis:level – typ studijního programu
- thesis:discipline – studijní program a studijní obor
- thesis:grantor – instituce přidávající titul

Prvky povinné pro Zprávy:

- nusl:resultID – označení grantu
- nusl:resultGrantGarant – poskytovatel projektu

Metadatový set EVSKP-MS v OAI-PMH

V případě záznamů VŠKP databáze VŠKP dle zákona o VŠ zpřístupněných prostřednictvím OAI-PMH serveru je doporučeno nad rámec povinného standardu Dublic Core (viz výše) použít Metadatový soubor pro elektronické vysokoškolské kvalifikační práce v ČR (EVSKP-MS). Tato metodika stanovuje, v souladu s EVSKP-MS, pro potřeby sdílení záznamů VŠKP protokolem OAI-PMH ve formátu EVSKP-MS tyto povinné prvky:

- dc:title - název VŠKP
- dc:creator – autor VŠKP
- dcterms:abstract – abstrakt VŠKP
- dcterms:dateAccepted – datum obhajoby VŠKP
- dc:type – typ VŠKP
- dcterms:medium – médium (formát souboru) VŠKP
- dc:identifier – identifikátor VŠKP
- dc:language – jazyk VŠKP
- thesis:degree – akademický titul nebo vědecko-pedagogická hodnost
- thesis:name – zkratka jména akademického titulu nebo vědecko-pedagogické hodnosti
- thesis:level – typ studijního programu
- thesis:discipline – studijní program a studijní obor
- thesis:grantor – instituce přidávající titul
- evskp:contact – identifikátor poskytovatele metadat

Doporučení:

Vždy myslet na tři hlavní pilíře: dodržet formát, dodržet protokol a mít konzistentní data.

Doporučujeme používat persistentní strategii mazání záznamů, protože je to jediný způsob, jak dát harvestujícímu repozitáři vědět, které záznamy byly smazány. Pokud by persistentní strategie použita nebyla, tak by se v harvestujícím repozitáři hromadily smazané neaktuální záznamy.

5.4 Předávání dat do národních a mezinárodních center

Digitální repozitáře šedé literatury podporují šíření výstupů vědecko-výzkumné práce, umožňují jednoduché sdílení výsledků mezi badateli a jsou hodnotným zdrojem informací pro vědu a vzdělávání. Pro zajištění lepší vyhledatelnosti a dostupnosti slouží repozitáře/agregátory sdružující data z jednotlivých institucionálních digitálních repozitářů. Pro zviditelnění výsledků práce instituce je nezbytné předávat data z digitálního repozitáře, do odpovídajících národních a mezinárodních repozitářů.

Co je potřeba splnit pro předání dat:

- 1) způsob předávání dat,
- 2) povinné prvky metadatového formátu,
- 3) mapování na druhy dokumentů,
- 4) další podmínky.

5.4.1 Způsob předávání dat

V dnešní době již většina repozitářů vyžaduje předávání dat pomocí OAI-PMH¹⁵ protokolu za účelem periodického, automatického a konzistentního sklizení dat. Podrobně je způsob předávání dat popsán v kapitole 5.3.2 Předávání dat.

5.4.2 Povinné prvky metadatového formátu

Předepsané metadatové formáty jsou různé, proto při předávání dat do národních a mezinárodních repozitářů je nutné dodržet požadovaná povinná pole konkrétního repozitáře.

Příklad povinných prvků metadatového formátu OpenGrey:

Snahou každého repozitáře šedé literatury by mělo být připojení se k mezinárodní databázi šedé literatury OpenGrey. Data mohou být odevzdávána prostřednictvím Digitálního repozitáře NUŠL, nebo přímo z konkrétního institucionálního repozitáře. Databáze OpenGrey využívá pro ukládání záznamů o dokumentech vlastní metadatový formát, který je tvořen prvky kvalifikovaného Dublin Core a vlastními specifickými prvky.

Při budování digitálního repozitáře je tak, v případě plánovaného odevzdávání dat do OpenGrey, potřeba zajistit implementaci povinných prvků.

Povinné prvky

Výčet prvků, které jsou povinné pro každý záznam. Jednotlivé prvky jsou zapsány dle předlohy (zvyklostí) OpenGrey.

¹⁵ <http://www.openarchives.org/OAI/openarchivesprotocol.html> [cit. 2010-10-10]

element=**“contributor”** qualifier=**“none”**

Bližší nespecifikovaná osoba, organizace nebo služba odpovědná za obsah zdroje.

element=**“country”** qualifier=**“iso”**

Kód země podle normy ISO 3166.

element=**“date”** qualifier=**“copyright”**

Rok vydání, podle normy ISO 8601 – pole se používá pro vyhledávání.

element=**“description”** qualifier=**“provenance”**

Zdroj záznamu (partnerská instituce OpenGrey).

element=**“language”** qualifier=**“iso”**

Jazyk dokumentu podle normy ISO 639-1.

element=**“subject”** qualifier=**“classification”**

Klasifikační kód třídění SIGLE.

element=**“title”** qualifier=**“none”**

Anglický název dokumentu (originální neb překlad názvu).

element=**“type”** qualifier=**“none”**

Typ dokumentu podle typologie OpenGrey.

Podmíněně povinné prvky

Prvky povinné pro některé případy, resp. pro specifické typy dokumentu.

element=**“contributor”** qualifier=**“advisor”**

Prvek především pro VŠKP, uvádí se vedoucí práce.

element=**“description”** qualifier=**“availability”**

Informace o dostupnosti dokumentu.

element=**“event”** qualifier=**“conference”**

Informace o konferenci zapsané podle pravidel OpenGrey (název, místo konání konference, datum konání konference).

element=**“event”** qualifier=**“date”**

Datum konání konference podle normy ISO 8101.

element=**“event”** qualifier=**“name”**

Název konference.

element=**“event”** qualifier=**“place”**

Místo konání konference, ve tvaru město (kód, nebo název státu).

element=**“partner”** qualifier=**“documentprovider”**

Instituce vlastníci dokument a kontakt na službu.

element=**“partner”** qualifier=**“documentproviderid”**

Odkaz na podrobné informace o poskytovateli dokumentu.

element="relation" qualifier="uri"

Link na plný text dokumentu.

element="thesis" qualifier="university"

Pole pro VŠKP, název univerzity.

5.4.3 Mapování na druhy dokumentů

Repozitáře a agregátory šedé literatury nemají společnou předepsanou typologii druhů dokumentů, ale zároveň je právě tento údaj vždy jedním z parametrů pro tvorbu sbírek a pro vyhledávání. Proto je nutné vždy namapovat mezi institucionálním repozitářem a národním či mezinárodním agregátorem odpovídající hodnoty a následně provést konverzi těchto dat buď na výstupu pro OAI-PMH set nebo až při vkládání. U mezinárodních repozitářů je většinou vždy požadována konverze již od dodávajícího institucionálního repozitáře.

5.4.4 Další podmínky

Agregátor si podle svého zaměření a funkcí může klást různé další podmínky. Například vlastní věcnou indexaci, zajištění služby dodání dokumentu apod. jako je tomu u evropské databáze šedé literatury OpenGrey.

5.4.5 Národní agregátory zaměřené na šedou literaturu

K posílení spolupráce na poli vědy, výzkumu a vzdělávání v České republice přispívají národní agregátory ověřené šedé literatury.

NUŠL

Národní úložiště šedé literatury zajišťuje jednotný přístup k šedé literatuře vzniklé v České republice. NUŠL je zaměřen na získávání metadat a plných textů. Velký důraz je přitom kladen na vytvoření partnerské sítě spolupracujících institucí z oblasti školství, výzkumu a vývoje, státní správy, komerčního sektoru a ověřených open access zdrojů.

www.nusl.cz

Theses.cz

„Theses.cz je systém pro odhalování plagiátů mezi závěrečnými pracemi a je vyvíjen a provozován Masarykovou univerzitou. Slouží vysokým školám a univerzitám (nejen v ČR) jako národní registr závěrečných prací (informací o pracích – název, autor, atd.) a jako úložiště prací pro vyhledávání plagiátů. Systém umožňuje zástupcům zapojených škol vkládat práce a vyhledávat mezi nimi plagiáty. Veřejnosti jsou zpřístupňovány záznamy o práci, příp. plné texty (dle rozhodnutí školy), a vyhledávání mezi nimi. Systém nabízí další služby, funkce a aplikace a bude rozvíjen dle potřeby.“¹⁶

<http://www.theses.cz/>

¹⁶<http://www.theses.cz/> [cit. 2010-05-10]

5.4.6 Mezinárodní agregátory zaměřené na šedou literaturu

Pro zviditelnění a dostupnost českého výzkumu a vědy je důležité předávat data také do mezinárodních repositářů/agregátorů. Pro připojení k agregátorům je vždy nutný OAI-PMH protokol pro předávání dat a dále splnit požadovaný metadatový formát včetně povinných polí, mapování na druhy dokumentů a případné další podmínky. Vzhledem k odlišným podmínkám u jednotlivých mezinárodních agregátorů je pro individuální instituce vhodné použít národního agregátora, který data zpracovává do požadovaných formátů a předává do mezinárodních repositářů/agregátorů.

DRIVER

DRIVER (Digital Repository Infrastructure Vision for European Research) je nadstavbou nad zapojenými evropskými institucionálními repositáři typu „open access“. Jedná se o vyhledávací portál, který vyhledává ve více jak 2,5 milionech článků, disertací, knih, zpráv apod. z 249 institucionálních a obsahově zaměřených repositářů z 33 zemí Evropy a v 25 jazycích.

<http://www.driver-repository.eu/>

Odkaz na DRIVER Guidelines

<http://www.techlib.cz/files/download/id/2476/driver-guidelines.pdf>

OpenGrey

Evropská databáze šedé literatury z oblasti vědy, techniky, ekonomie, humanitních a dalších věd. Databáze obsahuje 700 tisíc bibliografických záznamů technických zpráv, disertačních prací, konferenčních příspěvků a dalších druhů šedé literatury. OpenGrey má mimo jiné zcela specifické podmínky, např. nutnost dostupnosti digitálního textu nebo odkazu na službu, která dodá kopii z tištěné verze. Povinná je věcná indexace SIGLE. Upřednostňuje spolupráci přímo s národními agregátory.

<http://www.opengrey.eu/>

WorldWideScience.org

Projekt WorldWideScience.org je globální bránou vědy, která zajišťuje federativní vyhledávání v národních a mezinárodních vědeckých databázích a portálech s cílem zrychlit vědecký objev a pokrok.

<http://worldwidescience.org/>

Doporučení:

Pokud má instituce zájem předávat data do DRIVERu a OpenGrey, je pro ni nejvýhodnější předávat záznamy do NUŠL, který splňuje podmínky těchto agregátorů a postupně se připojuje k dalším.

Vzhledem k diskusím o autorskoprávní ochraně bibliografických dat a především k umožnění mezinárodní spolupráce doporučujeme bibliografická data explicitně označit jako volné dílo v režimu Open Bibliografic Data. Open Bibliografic Data je režim podporující otevřenost bibliografických dat. K vymezení práv k datům v open režimu je možné použít některou z mezinárodně platných licencí Public Domain Dedication and Location (PDDL) nebo Creative Commons Zero (CC0). Problematické volných licencí a výpis typů licencí jsou věnovány webové stránky Open Definition <http://opendefinition.org/okd/>. Toto doporučení je důležité zohlednit právě v lokálních repositářích, aby jasně určovali možnosti využití pro agregátory na všech úrovních.

6 Srovnání novosti postupů

Metodika zpracování, dlouhodobého uchování a zpřístupnění šedé literatury v České republice přináší zcela nové postupy a jako první popisuje zpracování šedé literatury v elektronické formě. Veškeré předchozí návody a postupy byly zaměřené pouze na tištěnou formu šedé literatury. Metodika nepřináší pouze nové postupy, ale poprvé přináší komplexní zmapování české legislativy související s šedou literaturou včetně doporučení a příkladů na implementaci do interních směrnic institucí. V celé metodice je kladen důraz na zpřístupnění digitálního obsahu a to jak na národní tak na mezinárodní úrovni. Proto je v metodice popsán postup předávání dat do národních a mezinárodních center. Zcela nově je v metodice zařazen také postup definování obsahu repozitáře, který je důležitý pro management institucí při určování náplně digitálního repozitáře.

7 Popis uplatnění certifikované metodiky

Metodika je určena pro samotné producenty šedé literatury, mezi které patří zejména vědecko-výzkumné organizace, vysoké školy, grantové agentury, orgány státní správy a samosprávy, paměťové instituce, ale také výrobní a obchodní společnosti. Metodika poskytuje cílové skupině postupy a legislativní rámec pro budování lokálních digitálních repozitářů tak, aby zajistili dlouhodobé uchování a zpřístupnění digitálních dat a aby jejich repozitáře splňovaly standardy potřebné pro národní i mezinárodní integraci.

Výsledky řešení jsou průběžně publikovány na odborných konferencích, seminářích v Česku i v zahraničí, ale také formou odborných článků a knihy v českém a anglickém jazyce. Veškeré výstupy a podpůrné materiály jsou prezentovány na webových stránkách projektu <http://nusi.techlib.cz>. Tyto stránky budou udržovány i po skončení projektu.

8 Použitá literatura

Archival Resources Key [online]. [cit. 2011-10-10]. Dostupné z www: <<https://wiki.ucop.edu/display/Curation/ARK>>.

Consultative Committee for Space Data Systems. 2002. *Reference model for an open archival information system (OAIS) : CCSDS 650.0-B-1* [online]. Washington (DC, USA) : CCSDS, January 2002 [cit. 2011-11-10]. Blue Book, Issue 1. Dostupný z WWW: <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>.

Creative Commons Česká republika. c2011. [cit. 2011-10-12]. Dostupný z www: <<http://www.creativecommons.cz/>>.

CUBR, Ladislav. *Dlouhodobá ochrana digitálních dokumentů*. Praha : Národní knihovna ČR, 2010. 154 s. ISBN 978-80-7050-588-5.

Digital-born [online]. Wikipedia, 2011. [cit. 2011-08-03]. Dostupné z www: <<http://en.wikipedia.org/wiki/Born-digital>>.

DOI [online]. DOI. ORG. Updated 4th October 2011. [cit. 2011-10-10]. Dostupné z www: <<http://www.doi.org/>>.

DRIVER [online]. Digital Repository Infrastructure Vision for European Research. [cit. 2011-10-10]. Dostupné z www: <<http://www.driver-community.eu/>>

DRIVER Guidelines [online]. [cit. 2011-10-10]. Dostupné z www: <<http://www.techlib.cz/files/download/id/2476/driver-guidelines.pdf>>

FRANTÍKOVÁ, B. Pokyny pro zpracování záznamů v systému CDS Invenio [online]. Praha : Národní technická knihovna. [2011]. Dostupné z www: <http://nusi.techlib.cz/images/Pokyny_zpracovani_zaznamu.pdf>.

FRANTÍKOVÁ, B.; PEJŠOVÁ P.. Manuál pro vkládání a editaci dat v systému CDS Invenio [online]. Praha : Národní technická knihovna. [2011]. Dostupné z www: <http://nusi.techlib.cz/images/Manual_Invenio1.pdf>.

Handle [online]. Corporation for National Research Initiatives 2011. cit. 2011-10-10]. Dostupné z www: <<http://www.handle.net/>>.

ISO 15836:2003 [online]. © ISO 2011. [cit. 2011-11-18]. Dostupné z www: <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37629>

MATUŠÍK, Z.. habilitační práce. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha: Národní knihovna ČR, 2003- [cit. 2011-10-11]. Dostupné z: <http://aleph.nkp.cz/F/?func=direct&doc_number=000000983&local_base=KTD>. (MATUŠÍK, 2003)

Národní technická knihovna [online]. c2006 – 2011. Aktualizováno 17. 6. 2011. [cit. 2011-10-12]. Dostupné z www: <www.techlib.cz>.

Národní úložiště šedé literatury [online]. Praha: Národní technická knihovna, 2008. Aktualizováno 23. 3. 2009, [cit. 2011-10-12]. Dostupné z www: <<http://nusi.techlib.cz/>>

Odborná komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací [online]. OK e-VŠKP AKVŠ ČR, c2004 – 2007. Aktualizováno 12. 10. 2007. [cit. 2010-03-12]. Dostupné z www: <<http://www.evskp.cz/>> .

OpenAIRE Guidelines 1.1 [online]. Listopad 2010. [cit. 2011-10-12]. Dostupné z www: <<http://www.openaire.eu/cs/component/attachments/download/79%E2%8C%A9=en>>.

Open bibliography and Open Bibliographic Data [online]. 2011 [cit. 2011-11-01]. Principy otevřených bibliografických dat. Dostupné z WWW: <<http://openbiblio.net/principles/cz/>>.

Open Definition [online]. 2011 [cit. 2011-11-01]. Defining the Open in Open Data, Open Content and Open Services. Dostupné z WWW: <<http://opendefinition.org/>>.

OpenGrey [online]. [cit. 2011-10-10]. Dostupné z www: <<http://www.opengrey.eu/>>

PEJŠOVÁ, Petra. Národní úložiště šedé literatury. Čtenář. 2010, roč. 62, č. 5, s. 176 - 180. ISSN 0011-2321 (PEJŠOVÁ, 2010a)

PEJŠOVÁ, P. (ed.). Repozitáře šedé literatury [online]. Zlín: VerBuM, 2010. s. 69. [cit. 2011-06-09]. Dostupný z www: <<http://nusi.techlib.cz/images/Book.pdf>>. ISBN 978-80-904273-5-8. (PEJŠOVÁ, 2010b)

POLČÁK, R.. Digitální zpracování tzv. šedé literatury pro Národní úložiště šedé literatury [online]. 2009. [cit. 2011-06-10]. Dostupné z www: http://nusi.techlib.cz/index.php/Právní_expertiza. (POLČÁK, 2009)

PURL [online]. OCLC. [cit. 2011-10-10]. Dostupné z www: <<http://purl.oclc.org/docs/index.html>>.

ŠÍMOVÁ, Marie. Identifikátory elektronických dokumentů: bakalářská práce. Praha: Vysoká škola ekonomická, 2010.

ŠVEJDA, J. referát. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha: Národní knihovna ČR, 2003- [cit. 2011-10-11]. Dostupné z: <http://aleph.nkp.cz/F/?func=direct&doc_number=000001770&local_base=KTD>. (ŠVEJDA, 2003)

The Open Archives Initiative Protocol for Metadata Harvesting [online]. [cit. 2011-10-10]. Dostupné z www: <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>.

TONER, D., Nad oprávněním zpracovávat osobní údaje [online]. 28. 10. 2003, [cit. 2011-06-09] Dostupné z: <<http://www.itpravo.cz/index.shtml?x=142872>>. (TONER, 2003)

Uniform Resource Name: National Bibliography Number [online]. [cit. 2011-10-10]. Dostupné z www: <<http://wiki.surffoundation.nl/display/standards/URN-NBN>>.

Vysokoškolské kvalifikační práce [online]. Fakulta informatiky Masarykovy univerzity. Aktualizováno 10. 6. 2010. [cit. 2010-05-10]. Dostupné z www: <<http://www.theses.cz/>>.

World wide science [online]. [cit. 2010-04-22]. World wide science aliance. Aktualizováno 10. 6. 2010. [cit. 2011-10-10]. Dostupné z www: <<http://worldwidescience.org/>>.

Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

Zákon č. 101/2000 Sb., zákon o ochraně osobních údajů a o změně některých zákonů (o ochraně osobních údajů).

Zákon č. 111/1998 Sb., zákon o vysokých školách a o změně a doplnění některých dalších zákonů (ZVŠ).

Zákon č. 257/2001 Sb. o knihovnách a podmínkách provozování veřejných knihovnických a informačních služeb (knihovní zákon).

Zákon č. 480/2004 Sb., o některých službách informační společnosti a o změně některých zákonů (zákon o některých službách informační společnosti).

Zákona č. 563/1991 Sb., o účetnictví (zákon o účetnictví).

Zákona č. 248/1995 Sb., o obecně prospěšných společnostech.

Zákona číslo 227/1997 Sb., o nadacích a nadačních fondech.

Zákona č. 106/1999 Sb., o svobodném přístupu k informacím.

9 Seznam publikací

Seznam publikací, které předcházely metodice a byly publikovány, případně výstupy z originální práce

FRANTÍKOVÁ, Bohdana. Šablony pro zpracování dat v systému CDS Invenio [online]. Praha : Národní technická knihovna. [2010]. Dostupné na WWW: <http://nysl.techlib.cz/images/Sablony_v1.pdf>.

FRANTÍKOVÁ, Bohdana. Pokyny pro zpracování záznamů v systému CDS Invenio [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nysl.techlib.cz/images/Pokyny_zpracovani_zaznamu.pdf>.

FRANTÍKOVÁ, Bohdana; SYNKOVÁ, Veronika; ŘÍMANOVÁ, Radka. Metadatový formát NUŠL. Verze 1.0 [online]. Praha : Národní technická knihovna. [2010]. 23 s. Dostupné na WWW: <http://nysl.techlib.cz/images/NUSLformat_verze1.0_1.pdf>.

FRANTÍKOVÁ, Bohdana. [Instrukce a informace pro instituce využívající CDS Invenio jako lokální repozitář](#) [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nysl.techlib.cz/index.php/Invenio_lokalni.pdf>.

FRANTÍKOVÁ, Bohdana. Instrukce a informace pro připojení k digitálnímu repozitáři NUŠL [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nysl.techlib.cz/images/Instrukce_pripojeni4.pdf>.

FRANTÍKOVÁ, Bohdana; PEJŠOVÁ, Petra. Manuál pro vkládání a editaci dat v systému CDS Invenio [online]. Praha : Národní technická knihovna. [2011]. Dostupné na WWW: <http://nysl.techlib.cz/images/Manual_Invenio1.pdf>.

FÜRSTOVÁ, Iveta. Projekt NUŠL a právní problematika šedé literatury. České právo a informační technologie. 9. -10. 9. 2010, Hrotovice, 2010. (v tisku)

FÜRSTOVÁ, Iveta. Šedá literatura a Národní úložiště šedé literatury. 11. konference Archivy, knihovny, muzea v digitálním světě 2010. Praha. 1. - 2. 12. 2010.

FÜRSTOVÁ, Iveta. Partnerská síť Národního úložiště šedé literatury (NUŠL). In 17. ročník konference Inforum, Vysoká škola ekonomická v Praze, Praha, 24. - 26. 5. 2011. Dostupný z WWW: <<http://www.inforum.cz/pdf/2011/furstova-iveta.pdf>>.

FÜRSTOVÁ, Iveta. Národní úložiště šedé literatury. Veřejná správa č. 13/2011, roč. XXII, vyšlo 30. 6. 2011, s. 18 – 19.

FÜRSTOVÁ, Iveta. Zařazení šedé literatury do Národního úložiště šedé literatury [online]. Praha : Národní technická knihovna. [2011], 18 s. Dostupné na WWW: <http://nysl.techlib.cz/images/Zarazeni_SL_do_NUSL_07_2011.pdf>.

FÜRSTOVÁ, Iveta; PEJŠOVÁ, Petra. Národní úložiště šedé literatury. In 16. ročník konference Inforum, Vysoká škola ekonomická v Praze, Praha, 25. - 27. 5. 2010. Dostupný z WWW: <<http://www.inforum.cz/pdf/2010/furstova-iveta-1.pdf>>.

MÜLLER, Tomáš. CDS Invenio - a software solution for National Repository of Grey Literature. In Seminar to access of grey literature 2010 : 3rd year of the seminar, 8. 12. 2010 [online]. Praha: Národní technická knihovna, 2010. Dostupný z WWW: <http://nysl.techlib.cz/index.php/Sbornik_2010>. ISSN 1803-6015.

Národní úložiště šedé literatury [online]. Praha : Národní technická knihovna, [2008]. Česká verze. Dostupná z WWW: <<http://nysl.techlib.cz/>>.

PEJŠOVÁ, Petra. Národní úložiště šedé literatury – digitální repozitář a centrální vyhledávací rozhraní. ITlib. Informačné technológie a knižnice. 2010, roč. 14, č. 2, s. 47 - 49. Dostupný z WWW: <<http://www.cvtisr.sk/itlib/itlib102/pejsova.htm>>.

PEJŠOVÁ, Petra. User Interface of the National Repository of Grey Literature. In Libraries in the digital age, University of Zadar, Zadar, Croatia, 24-28 May 2010.

PEJŠOVÁ, Petra (ed.). Repozitáře šedé literatury. Zlín : VeRBuM, 2010. 152 s. Dostupný z WWW: <<http://nysl.techlib.cz/images/Book.pdf>>. ISBN 978-80-904273-5-8.

PEJŠOVÁ, Petra. Czech National Repository od Grey Literature. The Grey Journal (TGJ): An international journal on grey literature / D.J. Farace (Journal Editor); J. Frantzen (Technical Editor); GreyNet, Grey Literature Network Service. - Amsterdam: TextRelease, Volume 6, Number 3, Autumn 2010, p. 117 - 127.

[PEJŠOVÁ, Petra. Projekt NUŠL a další projekty v ČR : Projekt „Digitální knihovna pro šedou literaturu – funkční model a pilotní realizace“. In Seminář ke zpřístupňování šedé literatury 2008 : 1. ročník semináře zaměřeného na problematiku uchovávání a zpřístupňování šedé literatury, 8. 10. 2008 \[online\]. Praha : Státní technická knihovna, 2008. Dostupný z WWW: <\[http://nysl.stk.cz/images/Projekt_NUSL_text.pdf\]\(http://nysl.stk.cz/images/Projekt_NUSL_text.pdf\)>. ISSN 1803-6015.](#)

PEJŠOVÁ, Petra. Projekt Digitální knihovna pro šedou literaturu - funkční model a pilotní realizace. Krok : Knižnická revue Olomouckého kraje. 2008, roč. 5, č. 4, s. 8 - 9. Dostupný z WWW: <<http://www.vkol.cz/data/soubory/import/krok0804.pdf>>.

[PEJŠOVÁ, Petra. System of Grey Literature in the Czech Republic. In Eleventh International Conference on Grey Literature, Library of Congress, Washington D.C., USA, 14-15 December 2009.](#)

[PEJŠOVÁ, Petra. National Repository of Grey Literature – second year. In Seminář ke zpřístupňování šedé literatury 2009 : 2. ročník semináře zaměřeného na problematiku uchovávání a zpřístupňování šedé literatury, Brno, 22. 10. 2009. Praha : Národní technická knihovna, 2009. Dostupný z WWW: <\[http://nysl.techlib.cz/images/NRGL_fulltext_2009_final.pdf\]\(http://nysl.techlib.cz/images/NRGL_fulltext_2009_final.pdf\)>. ISSN 1803-6015.](#)

PEJŠOVÁ, Petra. Národní úložiště šedé literatury jako součást Digitální polytematické knihovny vědy, výzkumu a vývoje. Knihovna : knihovnická revue. 2009, roč. 20, č. 1, s. 30 - 37. Dostupný z WWW: <<http://knihovna.nkp.cz/knihovna91/pejs.htm>>.

[PEJŠOVÁ, Petra. Projekt Digitální knihovna pro šedou literaturu po prvním roce řešení. Dostupný z WWW: <<http://www.schk.sk/fileadmin/docs/digilib/2009/Petra-PEJSOVA.pdf>>.](#)

PEJŠOVÁ, Petra. Národní úložiště šedé literatury. Blok expertů, Filozofická fakulta Masarykovy univerzity. Brno. 16. 12. 2010.

PEJŠOVÁ, Petra. National Repository of Grey Literature - third year. In Seminar to access of grey literature 2010 : 3rd year of the seminar, 8. 12. 2010 [online]. Praha : Národní technická knihovna, 2010. Dostupný z WWW: <http://nysl.techlib.cz/index.php/Sbornik_2010>. ISSN 1803-6015.

PEJŠOVÁ, Petra. Národní úložiště šedé literatury (NUŠL). Čtenář : měsíčník pro knihovny. 2010, roč. 62, č. 5, s. 176-180. Dostupný z WWW: <<http://ctenar.svkkk.cz/clanky/2010-roc-62/05-2010.htm>>.

PEJŠOVÁ, Petra. Audit důvěryhodnosti digitálního repozitáře NUŠL. Digitální knižnice 2011, Jasná pod Chopkom, 28. – 30. 3. 2011.

PEJŠOVÁ, Petra; FÜRSTOVÁ, Iveta. Uživatelské rozhraní Národního úložiště šedé literatury. Digitální knižnice 2010, Jasná pod Chopkom, 22. – 24. 9. 2010.

[PEJŠOVÁ, Petra; PFEIFEROVÁ, Martina. Grey Literature in the Czech Republic. In Tenth International Conference on Grey Literature : Designing the Grey Grid for Information Society : Amsterdam, Netherlands, 8-9 December 2008 \[online\].](#) 1st ed. Amsterdam : TextRelease, 2009. ISBN 978-90-77484-11-1. ISSN 1386-231. Dostupný z WWW: <<http://opensigle.inist.fr/handle/10068/697786>>.

PEJŠOVÁ, Petra; POLČÁK, Radim. Právní aspekty šedé literatury. In Seminář Právní aspekty e-learningu. Právnická fakulta Masarykovy univerzity, Brno. 16. 6. 2010.

PEJŠOVÁ, Petra; VASKA, Marcus: An Analysis of Current Grey Literature Document Typology. In Twelfth International Conference on Grey Literature, National Technical Library, Prague, Czech republic, 6-7 December 2010.

Persistentní identifikátory pro NUŠL – rozhodovací kritéria [online]. Praha : Národní technická knihovna. [2009]. Dostupné na WWW: <http://nusl.techlib.cz/images/PID_kriteria.pdf>.

POLČÁK, Radim. Digitální zpracování tzv. šedé literatury pro Národní úložiště šedé literatury [online]. Brno: Právnická fakulta Masarykova univerzita, 2009. Dostupné z WWW: <http://nusl.techlib.cz/images/Pravni_exp_NUSL.pdf>.

Specifikace SW pro NUŠL [online]. Praha : Národní technická knihovna. [2008]. Dostupné na WWW: <http://nusl.techlib.cz/images/SW_specif.pdf>.

Typologie Národního úložiště šedé literatury [online]. Praha : Národní technická knihovna. [2008]. Dostupné na WWW: <http://nusl.techlib.cz/images/Typologie_celek_text.pdf>.

10 Seznam zkratek

A

AMU - Akademie múzických umění v Praze
ASCII - American Standard Code for Information Interchange
AVI – Audio Video Interleave
ASEP - Automatizovaný systém evidence publikací
ARK - Archival Resource Key

B

Bc. - bakalář

C

CC – Licence Creative Commons
CCO - Creative Commons Zero

D

DCMI - Abstract Model
DOI - Digital Object Identifier, Identifikátor digitálního objektu
DRIVER - Digital Repository Infrastructure Vision for European Research

E

EDT-MS - standard pro popis eVŠKP
eVŠKP - elektronické vysokoškolské kvalifikační práce
EVSKP-MS - metadatový soubor pro eVŠKP v České republice

F

FTP - File Transfer Protocol (protokol pro přenos souborů)

G

GL – International Conference on Grey Literature

H

HTML – Hypertext Markup Language

I

Ing. – inženýr
IP (adresa) – internetový protokol
ISBN - International Standard Book Number, mezinárodní standardní číslo knihy
ISSN - International Standard Serial Number, identifikuje periodika
ISO - International Standard Organization; Mezinárodní organizace pro standardizaci
IS VaV – Informační systém výzkumu a vývoje

J

JPG – Point Photographic Experts Group (metoda ztrátové komprese digitálních obrázků)
JUDr. – doktor práv

L

LDAP - Lightweight Directory Access Protocol

M

Mgr. – magistr

MP3 –MPEG-2 Audio Layer III (metoda ztrátové komprese zvukových souborů)

MP4 – MPEG-4 Part 14 (multimediální kontejner)

MPEG – Motion Picture Experts Group (metoda ztrátové komprese videa a zvuku)

MS DOC – Microsoft Document

MS DOCx – Microsoft Dokument Extended

MS PPT – Microsoft Powerpoint

MS PPTx – Microsoft Powerpoint Extended

MŠMT – Ministerstvo školství, mládeže a tělovýchovy ČR

MUDr. – doktor medicíny

MVDr. – doktor veterinární medicíny

N

NTK - Národní technická knihovna, do 1. 7. 2009 Státní technická knihovna (STK)

NUŠL - Národní úložiště šedé literatury

O

OAI - Open Archives Initiative

OAI-PMH - Open Archives Initiative Protocol for Metadata Harvesting

OAIS - Open Archival Information System

P

PDDL - Public Domain Dedication and Location

Ph.D. - doktor

PhDr. – doktor filozofie

PID – Persistentní identifikátor

PDF - Adobe Portable Document Format

PNG – Portable Network Graphics

PSH - Polytematický strukturovaný heslář

PURL - Persistent Uniform Resource Locators

R

RBAC - Role Based Access control

RNDr. – doktor přírodních věd

RSS - Really Simple Syndication; je soubor s příponou .xml a je psán v jazyce XML

S

SRU – Search/Retrieve via URL

T

TDKIV - Česká terminologická databáze knihovnictví a informační vědy

ThDr. – doktor teologie

Theses.cz - Národní registr VŠKP a systém na odhalování plagiátů

ThLic. – licenciát teologie

TEF 2.0 - Thèses Electroniques Françaises

TIFF - Tag Image File Format

TXT – Text file

U

UNICODE – Tabulka znaků všech abeced

ÚOOÚ – Úřad pro ochranu osobních údajů

URI - Uniform Resource Identifier

URN:NBN - Uniform Resource Name: National Bibliography Number

URL - Uniform Resource Locator

UTC - Coordinated Universal Time (koordinovaný světový čas)

UTF-8 – UCS Transformation Format (způsob kódování textových řetězců)

V

VaV - Výzkum a vývoj v České republice

VaVaI – Výzkum, vývoj a inovace

VŠKP - vysokoškolské kvalifikační práce

v.v.i. – vědecká výzkumní instituce

W

www - informační systém pro práci s hypertextovými dokumenty na Internetu (angl. World Wide Web, "celosvětová pavučina")

X

XML - eXtensible Markup Language; rozšiřitelný značkovací jazyk

Z

ZOOÚ – Zákon na ochranu osobních údajů

ZVŠ – Zákon o vysokých školách

Dedikace:

Program DC - Zpřístupnění a ochrana kulturních, uměleckých a vědeckých zdrojů (2008-2011) financovaného MK ČR pod identifikačním kódem DC08P02OUK007 „Digitální knihovna pro šedou literaturu – funkční model a pilotní realizace“

Oponenti:

- 1) Ing. Petr Žabička, Moravská zemská knihovna v Brně
- 2) Mgr. Andrea Fojtů, Univerzita Karlova

Vydala: Národní technická knihovna, Praha, 2011