

# ŠEDÁ LITERATURA V STK: NOVÉ AKTIVITY, NOVÉ OBZORY

*Petr Novák, Státní technická knihovna*

## Úvod – historie šedé literatury ve Státní technické knihovně

### **SIGLE, EAGLE a STK**

Počátkem 90. let 20. století zástupci **Evropské asociace pro využívání šedé literatury EAGLE** (European Association for Grey Literature Exploitation) vyzvali **Státní technickou knihovnu (STK)** a **Knihovnu Akademie věd ČR (KNAV)** k účasti v asociaci. Záměrem bylo bibliograficky podchytil produkci šedé literatury v ČR a zpřístupnit ji uživatelům **Evropského systému informací pro šedou literaturu SIGLE** (System of Information for Grey Literature in Europe). Výsledkem jednání byl vstup STK do EAGLE v roce 1995. Již s předstihem došlo v průběhu roku 1994 k vytvoření **Kooperačního systému šedé literatury (KSŠL)** v STK. Tento systém, založený na smluvním základě, přebíral bibliografické záznamy o šedé literatuře (zejména o disertacích) od kooperujících českých vysokých škol, převzatá metadata konvertoval do specifického datového formátu SIGLE a v dávkách zasílal do FIZ Karlsruhe, kde byla data zpracovávána. STK současně zpracovávala pro SIGLE dokumenty šedé literatury, které měla přímo ve svém fondu.

Dlužno dodat, že spolupracující vysoké školy měly možnost předávaná data předvybrat dle relevance. Zdaleka tedy nešlo o úplnou evidenci produkce disertací v ČR. Role Knihovny AV ČR, která měla podchycovat šedou literaturu z produkce ústavů AV ČR nebyla za celou dobu existence systému SIGLE naplněna.

**Šedá literatura** – možné definice (sestavila Štěpánka Žižková)

1. informační zdroje vydávané prostřednictvím neoficiálních a neformálních publikačních cest. Je nekonvenční nebo polopublikovaná – tj. šedá (PhDr. Marta Mandysová, Slovenská technická knižnica);

2. dokumenty, které nejsou lehce přístupné prostřednictvím tradičních knihkupeckých kanálů a proto se těžko vyhledávají a získávají (British Library);

3. napsané nebo vytištěné dokumenty, určené omezené skupině uživatelů mimo okruh komerčního vydávání a rozšiřování a na rozhraní prostředků bibliografické kontroly (Francouzská asociace pro normalizaci – AFNOR);

4. polopublikované dokumenty, které nejsou publikovány obvyklým způsobem a nejsou proto dostupné na běžném knižním trhu (např. diplomové a disertační práce, výzkumné zprávy, interní dokumenty, oficiální publikace atd.). Pro vyhledávání a distribuci šedé literatury existují specializované informační systémy (např. databáze SIGLE) (TDKIV).

**Typologie šedé literatury:** výzkumné zprávy, vysokoškolské kvalifikační práce, konferenční materiály, sborníky, cestovní zprávy, firemní literatura, interní dokumenty organizací.

**Charakteristika šedé literatury:**

1. není běžně dostupná prostřednictvím knihkupectví a dalších oficiálních kanálů;

2. nesnadno se vyhledává;

3. nesnadno se získává;

4. specifická pravidla zpřístupňování úplných textů;

5. přináší nejčerstvější poznatky, formuje state-of-art všech oborů vědy a výzkumu.

Zakládající členské státy asociace EAGLE (především Velká Británie, Německo, Francie, Itálie, a Nizozemsko) do systému SIGLE přispívaly řádově větším objemem včetně dalších druhů šedé literatury (výzkumné a technické zprávy apod.). Ze zemí bývalého východního bloku byla Česká republika prakticky jediným státem, který se systematicky a v přijatelném rozsahu snažil o podchycení produkce šedé literatury a o zpřístupnění bibliografických dat uživatelům z ostatních států. Ostatní východoevropské země se na činnosti EAGLE/SIGLE podílely pouze okrajově. Financování systému SIGLE záviselo na členských příspěvcích jednotlivých členských států a na výnosech prodeje databáze SIGLE na CD-ROM (SilverPlatter) a zpřístupnění této databáze v databázovém centru FIZ Karlsruhe. Systém na jedné straně disponoval relativně propracovaným systémem věcného třídění zvaným Seznam předmětových kategorií SIGLE (SIGLE Subject Category List), na druhé straně byl omezen metadatovým formátem, který byl nejen nesnadno rozšiřitelný a zcela nekompatibilní s kterýmkoliv z formátů rodi-

ny MARC, ale zejména zcela postrádal přímou vazbu umožňující snadné získání plného textu dokumentu. Západoevropské členské státy tento nedostatek řešily pomocí služeb dodávání dokumentů (document delivery) mateřských knihoven. Tato infrastruktura však v ČR neexistovala a proto se požadavky na výpůjčky z ČR, případně do ČR řešily cestou standardní MVS bez jakékoli zjišitelné vazby na systém KSSL/SIGLE.

V dalších letech se finanční otázky staly jedním z omezujících faktorů systému, jehož omezený rozpočet neumožňoval jeho rozšiřování o moderní služby a aplikace; podobně i formát SIGLE nedokázal udržet krok s dynamicky se měnícím prostředím charakterizovaným nástupem Internetu coby přenosového média i publikačního prostoru. Personální kapacita EAGLE nepostačovala, k čemuž navíc přispěly i v rychlém sledu jdoucí personální změny ve vedení asociace, zejména její Technické komise. Uvažované způsoby modernizace (rozšíření počtu členů, dopracování vazeb na služby dodávání dokumentů) nakonec nebyly realizovány právě z výše uvedených důvodů.

Celková koncepční a technologická zastaralost systému SIGLE, který přestal odpovídat nárokům kladeným na moderní prostředek shromažďování údajů o šedé literatuře, ale zejména nárokům na její bezbariérové zpřístupnění, vedla Velkou Británii, Nizozemsko, klíčové členy EAGLE, v letech 2003–2004 k vystoupení z asociace a zastavení spolupráce v systému SIGLE. Stejný krok ohlásily Německo a Francie. Systém resp. jeho klíčový produkt, databáze SIGLE, se tak stala nereprezentativní, počet nových záznamů drasticky poklesl a produkt se tak stal nevyužitelným a neschopným konkurence. EAGLE byla proto v roce 2005 rozpuštěna. Následující pokusy o resuscitaci systému v podobě systému MetaGrey iniciované Francií se (zatím) nesetkaly s úspěchem.

Státní technická knihovna za dobu provozu KSSL do systému SIGLE přispěla cca 5500 záznamy disertací (zlomek produkce ČR) s ročním přírůstkem v řádu stovek záznamů. Příčinou nízkého počtu byly jednak poměrně striktní požadavky SIGLE na jazykové vybavení dokumentu i záznamu, jednak zcela dobrovolná účast českých účastníků systému.

### ***Šedá literatura v STK v letech 2005–2006: druhé dějství?***

Bezprostředně po ukončení účasti v EAGLE vedení STK posoudilo možnosti pokračování ve svých aktivitách v oblasti šedé literatury s předpokladem, že STK bude i nadále schopna zpracovávat disertační práce, ale i další typy šedé literatury, které v budoucnu získá do svého fondu v klasické tištěné podobě.

V roce 2005 proběhlo několik jednání a setkání odborníků, iniciovaných STK, na kterých byly zkoumány možnosti a směry spolupráce

v oblasti šedé literatury v ČR. Na schůzce zástupců vysokých škol, ústředních knihoven a dalších institucí v STK v červnu 2005 byla **deklarována vůle k další spolupráci v této oblasti**. STK byla pověřena započít práce na vytvoření „národního centra pro šedou literaturu“. Jako možný či nosný směr dalšího vývoje STK považuje ve sblížení svých aktivit s činností **Oborné komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací AKVŠ** [1]. Tato komise, reprezentující skupinu producentů jednoho z nejdůležitějších a nejhodnotnějších typů šedé literatury, si stanovila cíle, které jsou konformní se záměry STK.

STK proto nabídla své více než desetileté zkušenosti, technologie a kapacitu k vývoji životaschopného a moderního nástroje, umožňujícího podchytit data o VŠKP<sup>1</sup> a dalších typech šedé literatury, jejich zpracování a zpřístupnění pokud možno v elektronické podobě.

Tento příspěvek shrnuje stav aktivit v dané oblasti k polovině roku 2006.

### **Současnost šedé literatury v ČR a ve světě**

Otázky spojené s tvorbou, šířením, zpracováním a využíváním dokumentů šedé literatury se díky neustálým technologickým inovacím mění. Tato kapitola se zabývá významnými světovými i domácími dokumenty týkajícími se této problematiky a vypočítává zdroje šedé literatury v ČR včetně definice právního rámce týkajícího se významné složky šedé literatury, totiž vysokoškolských kvalifikačních prací.

### ***Současná vědecká komunikace: aktuální trendy***

Koncepce shrnuté v této kapitole charakterizují změny v chápání komunikování vědeckých informací a změny úlohy knihoven v prostředí elektronické komunikace, Internetu a digitálních technologií. Knihovny se z pasivního zpracovatele a jakési půjčovny produkce vědeckých dokumentů ve stále rostoucí míře stávají aktivním spoluúčastníkem – nebo alespoň technickým podporovatelem – vytváření, zpřístupňování, distribuce a dlouhodobého uchování dokumentů produkovaných v mateřské instituci. Jedním ze způsobů iniciace této změny je provozování specifických digitálních knihoven – repositářů neboli úložišť.

### **Studie OCLC**

Změny prostředí komunikování vědeckých informací a úlohu paměťových institucí nastiňuje zpráva OCLC z roku 2003 nazvaná The OCLC Environmental Scan [2]:

---

<sup>1</sup> VŠKP – vysokoškolské kvalifikační práce všech stupňů (bakalářské, diplomové, disertační, rigorózní, habilitační). Jde o jednu z klíčových součástí šedé literatury.

*„Nové toky vědeckých informací jsou formovány od dob rozšíření Internetu, od dob vytvoření alternativních publikačních modelů, od dob systematického využívání Internetu knihovnami, od dob vzniku nových typů ukládání dokumentů. Doba dynamických proměn stále trvá a my jsme jejími přímými účastníky, diváky i iniciátory. Přestože tradiční modely s jasnými obrysy přetrvávají, v pozadí, v podzemí a za oponou vystávají systémy, přístupy a zvyklosti, které lze označit jako „under construction“. Následující schémata nastiňují proměny prostředí, kterých jsme svědky:*

**Repozitáře<sup>2</sup>** – institucionální, osobní nebo komunitní, oborově i obecně zaměřené.

**Služby agregátorů** – může se jednat o komerční služby (EBSCO, ProQuest) nebo služby vytvářené uživatelskými obcemi (ResearchIndex).

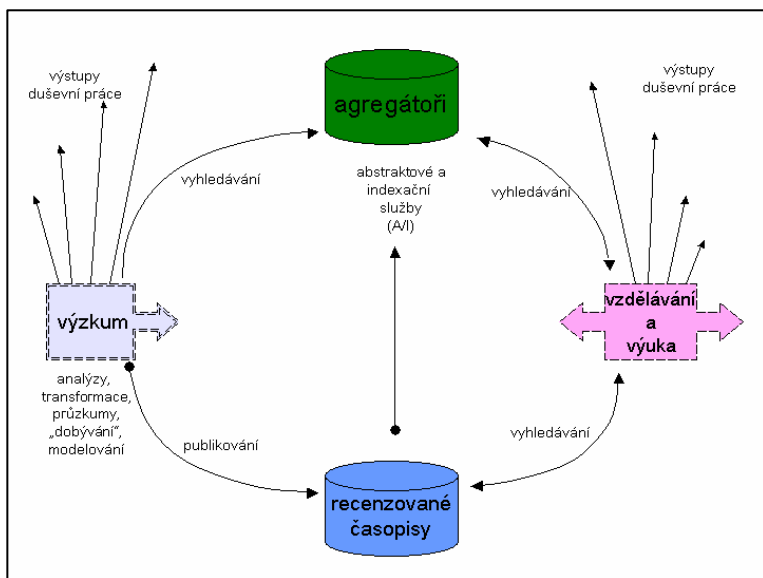
**Knihovny** – jejich potenciál v oblasti šíření informací novými kanály je nezpochybnitelný. Role knihoven může nabývat různých rozměrů:

- správce institucionálních repositářů,
- držitel/správce licence komerčních služeb poskytovaných externími dodavateli,
- prostředník pro ukládání a samoobslužné archivování,
- poskytovatel služeb lokálních trvalých repositářů.

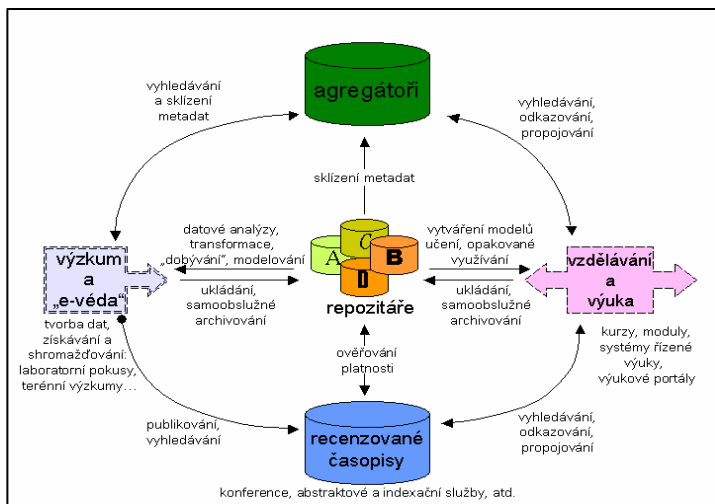
*Knihovny díky své podstatě informačních a paměťových institucí mají možnost a schopnost určovat budoucí vývoj v oblasti systémů správy znalostí. Jsou předurčeny k tomu, aby pokračovaly ve vytváření vazeb a vztahů, zprostředkovaly služby umocňující hodnoty vytvářené vědeckými obcemi a potvrdily svoji roli důvěryhodných institucí pro oblast vědy, výzkumu a vzdělávání.*

---

<sup>2</sup> Termín repositář (repository, repositorium) má několik významů: studnice (např. také moudrosti), spížirna, skladiště, archiv, regál lékárny. Termín lze přeložit do češtiny také jako úložiště, čehož využili autoři názvu projektu STK.



Obr.1 - Tok vědeckých informací – tradiční model [2]



Obr. 2 - Tok vědeckých informací dnes [2]

## **i2010 : Digitální knihovny**

Problematiku digitalizace a zpřístupňování digitálních zdrojů si uvědomuje i Evropská unie, která ve „Sdělení Evropské komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů“ nazvaném „i2010 : Digitální knihovny“ [3] definuje úlohu knihoven takto:

*„Uchovávání digitálních záznamů je pro informační společnost, kde nabídka informací exponenciálně roste a obsah se stává stále více dynamický, životně důležitým problémem. V současnosti jsou s uchováváním digitálních záznamů malé zkušenosti, právní rámec se vyvíjí, prostředky jsou omezené a výsledek úsilí o uchování je nejistý. Problém si zasluhuje zvýšenou pozornost politiků i nejvíce zainteresovaných subjektů. Jeho dopad dalece přesahuje oblast knihoven a archivů a týká se všech organizací, které vytvářejí digitální informace a mají zájem na zachování jejich dostupnosti.“*

## **Studie ekonomického a technického rozvoje trhu vědeckého publikování v Evropě**

Důležitost digitálních úložišť zmiňuje „Studie ekonomického a technického rozvoje trhu vědeckého publikování v Evropě“ [4] z roku 2006, zpracovaná Evropskou komisí. V kapitole nazvané „Přístup k výsledkům výzkumu“ se píše:

*„Vzhledem k rozpočtovým problémům knihoven, příležitostem, které nabízejí moderní technologie, a se zřetelem k tomu, jak významná část veřejných prostředků se podílí na procesu vědeckého publikování, začínají se vědecké komunity a organizace na podporu výzkumu ve zvýšené míře zabývat o problematiku vědecké komunikace, a to s ohlasem na straně občanské společnosti a politické reprezentace na národní i mezinárodní úrovni.“*

*Roste zájem o to, aby za účelem co největší podpory výzkumu byly jeho výsledky obecně šiřitelné a volně dostupné. Volně dostupná literatura (open-access literature) se vyznačuje tím, že je digitální, on-line, bezplatná, a osvobozená od většiny autorskoprávních a licenčních omezení.“*

*Kromě možnosti zveřejnění na vlastních webových stránkách autorů se naskýtají dva základní způsoby, kterých mohou autoři článků využívat: (1) ukládat dokumenty ve volně dostupných archivech a úložištích, anebo (2) publikovat ve volně dostupných recenzovaných odborných periodikách.“*

V závěru studie je uvedena řada doporučení, mezi nimi i týkající se rozvoje zmíněných modelů publikování:

- *přístup veřejnosti k výsledkům krátce po publikování,*
- *stálá dostupnost,*
- *interoperabilita s dalšími systémy umožňující další šíření,*

- *další výzkum, zejména v oblasti autorského práva, alternativních modelů šíření a vývoje souvisejících technologií.*

### **Dlouhodobý záměr MŠMT pro oblast vysokých škol na období 2006 - 2010**

Roli vysokoškolských knihoven ve sféře tvorby infrastruktury pro výzkum a vývoj definovalo Ministerstvo školství, mládeže a tělovýchovy ve svém „*Dlouhodobém záměru vzdělávací a vědecké, výzkumné, vývojové, umělecké a další tvůrčí činnosti pro oblast vysokých škol na období 2006 – 2010*“ [5] takto:

*„Vysokoškolské knihovny budou zabezpečovat tvorbu informačních fondů, jejich dostupnost a dostupnost dalších informačních zdrojů pro potřeby VaV na vysokých školách, zprostředkovávat jejich maximálně efektivní využití a přispívat k vyhodnocování, šíření a aplikaci výsledků VaV a využívání výsledků diplomových a disertačních, příp. i dalších prací“.*

### **Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010**

„Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010“ [6], v současnosti připravovaná Národní knihovnou a Ministerstvem kultury jako dílčí cíl „Koncepce rozvoje knihoven v České republice na léta 2004 až 2010“, problematiku šedé literatury či zdrojů pro vědu a výzkum přímo nezmiňuje, nicméně definuje odpovědnost institucí za zdroje mimo jádro národního kulturního dědictví takto:

*„Širší Česká digitální knihovna je však tvořena velkým množstvím dalších digitálních dokumentů oborového, regionálního, institucionálního i jiného charakteru. Pouze některé z těchto zdrojů se kvalifikují jako nejcennější součást národního kulturního dědictví. Za shromažďování, trvalé uchovávání (na centrálním datovém úložišti nebo lokálních datových úložištích) i zpřístupnění zdrojů mimo "jádro" národního kulturního dědictví nesou odpovědnost (včetně finanční) resortně příslušná ministerstva, regiony, instituce atd. V rámci jejich působnosti existují dílčí programy a projekty, které je třeba efektivně propojit mezi sebou i s jádrem. Tento model do značné míry kopíruje osvědčený model shromažďování, trvalého uchovávání a zpřístupňování tradičních dokumentů.“*

### **Zdroje šedé literatury**

Jakých dokumentů, institucí a zdrojů se týká problematika současné šedé literatury? Zdroje lze shrnout do tří základních okruhů.



První okruh představují vysoké školy, realizující rozsáhlou publikační činnost v celém spektru svých specializací. Publikační činnost zahrnuje jak aktivity samotných pracovníků (studijní materiály, příspěvky do sborníků), tak výsledky vědecké a výzkumné práce jednotlivých vysokoškolských pracovišť v podobě technických a výzkumných zpráv. Nezastupitelnou roli zde mají vysokoškolské kvalifikační práce všech typů.

Druhý okruh lze obdobně vymezit publikační činností pracovníků Akademie věd ČR, evidovanou v systému ASEP (Automatizovaný systém evidence publikací).

Třetí okruh je vztažen k pravidelně či nepravidelně pořádaným aktivitám typu konferencí, výstav či workshopů, které publikují konferenční sborníky, katalogy, programy apod.

Doplňkovou roli lze spatřovat jak v již existujících sbírkách šedé literatury v knihovnách, zejména ve fondu Státní technické knihovny (disertace), tak i v publikačních aktivitách dalších výzkumných organizací, jako jsou podnikové či soukromé výzkumné ústavy. V neposlední řadě do této sféry spadají aktivity pracovníků institucí státní správy a samosprávy (ministerstva, kraje, státem zřízené organizace). Do sféry šedé literatury zcela jistě patří i firemní literatura jako cenný zdroj informací z praxe.

Forma dokumentů šedé literatury může být jak charakteru klasického tištěného média, tak elektronického dokumentu dostupného online (na Intranetu či v uzavřené LAN) nebo na offline nosiči typu CD/DVD.

### ***Právní otázky***

Právní rámec pro zveřejňování vysokoškolských kvalifikačních prací definuje **novela zákona o vysokých školách 552/2005 Sb** [7]. v paragrafu §47b takto:

§ 47b - Zveřejňování závěrečných prací

(1) Vysoká škola nevýdělečně zveřejňuje disertační, diplomové, bakalářské a rigorózní práce, u kterých proběhla obhajoba, včetně posudků oponentů a výsledku obhajoby prostřednictvím databáze kvalifikačních prací, kterou spravuje. Způsob zveřejnění stanoví vnitřní předpis vysoké školy.

(2) Disertační, diplomové, bakalářské a rigorózní práce odevzdané uchazečem k obhajobě musí být též nejméně pět pracovních dnů před konáním obhajoby zveřejněny k nahlížení veřejnosti v místě určeném vnitřním předpisem vysoké školy nebo není-li tak určeno, v místě pracoviště vysoké školy, kde se má konat obhajoba práce. Každý si může ze zveřejněné práce pořizovat na své náklady výpisy, opisy nebo rozmnoženiny.

(3) Platí, že odevzdáním práce autor souhlasí se zveřejněním své práce podle tohoto zákona, bez ohledu na výsledek obhajoby.“

Kromě vysokoškolského zákona se problematiky týká **autorský zákon** (novela 216/2006Sb.) [8].

V § 30 je definováno volné užití

(1) Za užití díla podle tohoto zákona se nepovažuje užití pro osobní potřebu fyzické osoby, jehož účelem není dosažení přímého nebo nepřímého hospodářského nebo obchodního prospěchu, nestanoví-li tento zákon jinak.

(2) Do práva autorského tak nezasahuje ten, kdo pro svou osobní potřebu zhotoví záznam, rozmnoženinu nebo napodobeninu díla.

(3) Nestanoví-li tento zákon dále jinak, užitím podle tohoto zákona je užití počítačového programu či elektronické databáze i pro osobní potřebu fyzické osoby či vlastní vnitřní potřebu právnické osoby nebo podnikající fyzické osoby včetně zhotovení rozmnoženiny takových děl i pro takovou potřebu; (...)

Práva knihoven a obdobných nevýdělečných zařízení školského a vzdělávacího charakteru (archivy, muzea, galerie, školy včetně vysokých) stanovuje § 37:

(1) Do práva autorského nezasahuje knihovna, archiv, muzeum, galerie, škola, vysoká škola a jiné nevýdělečné školské a vzdělávací zařízení,

a) zhotoví-li rozmnoženinu díla, která neslouží k přímému nebo nepřímému hospodářskému nebo obchodnímu účelu, pro své archivní a konzervační potřeby,

b) zhotoví-li rozmnoženinu díla, jehož rozmnoženina byla poškozena nebo ztracena, a o němž lze na základě rozumně vynaloženého úsilí zjistit, že není nabízeno k prodeji, nebo tiskovou rozmnoženinu malé části díla, jež byla poškozena nebo ztracena; takovou oprávněně zhotovenou rozmnoženinu může také podle odstavce 2 půjčovat,

c) zpřístupňuje-li dílo, včetně zhotovení jeho rozmnoženiny nezbytné pro takové zpřístupnění, které je součástí jeho sbírek a jehož užití není předmětem prodejních nebo licenčních podmínek, s výjimkou sdělování díla způsobem uvedeným v § 18 odst. 2, jednotlivcům ze strany veřejnosti prostřednictvím k tomu určených technických zařízení umístěných v jeho objektech, a to výhradně pro účely výzkumu nebo soukromého studia takových osob, a zamezí-li takovým osobám zhotovit rozmnoženinu díla; ustanovení § 30a odst. 1 písm. c) a d) tím nejsou dotčena,

d) půjčuje-li originály nebo rozmnoženiny obhájených diplomových, rigorózních, disertačních a habilitačních prací na místě samém, a to výhradně pro účely výzkumu nebo soukromého studia, pokud takové užití autor nevyloučil.

## **Projekt Národního úložiště šedé literatury (NUŠL)**

### ***Idea Národního úložiště ŠL<sup>3</sup>***

Jedním z hlavních úkolů Státní technické knihovny na rok 2006 je „Zpracování koncepce trvalého uchování a zpřístupnění národní produkce šedé literatury“. V jeho rámci je zpracovávána koncepce metadatového systému, jež vychází z koncepce dílčích distribuovaných repozitářů šedé literatury vytvářených na vysokých školách, v AV ČR a dalších institucích, poskytne integrovaný/souborný přehled o podchycených dokumentech šedé literatury a umožní jejich zpřístupnění.

Při formulaci tohoto úkolu vycházela STK ze své role ústřední vědeckotechnické knihovny pro oblast školství, vědy a výzkumu. STK je gestorem činností směřujících k podchycení, uchovávání a zpřístupňování dokumentů šedé literatury.

STK si je vědoma, že současný stav nepodporuje přístup k plným textům ŠL, brání účinnému šíření vědeckých poznatků a zkresluje pohled na vědeckovýzkumný potenciál ČR jako celku.

Předpokládá se, že Národní úložiště šedé literatury (NUŠL) bude „institucionalizovaným“ vyjádřením aktivit týkajících se šedé literatury v projektované Národní technické knihovně (NTK) a bude tvořit jednu z jejích hlavních činností. Vzhledem ke skutečnosti, že NTK vznikne sloučením dvou významných vysokoškolských knihoven (ČVUT a VŠCHT) a STK, lze předpokládat významné zapojení obou vysokoškolských subjektů do přípravy a vývoje NUŠL.

### ***Východiska projektu NUŠL***

#### **NUŠL a Koncepce trvalého uchování knihovních sbírek**

*Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010* [6] (viz též kapitola 2.1.5) je v současnosti připravována NK ČR jako dílčí cíl *Koncepce rozvoje knihoven v České republice na léta 2004 až 2010* (schválené usnesením vlády ČR č.679 ze dne 7. července 2004). V základním funkčním schématu této koncepce se prostor pro NUŠL nachází v tzv. standardní (fialové) zóně, která předpokládá zastřešení prostřednictvím portálů a souborných katalogů digitálních dokumentů. NUŠL by bylo jedním z nich. Předpokládá se přitom, že zodpovědnost za tato úložiště (včetně finanční) ponесou jednotlivá resortně příslušná ministerstva (v případě NUŠL tedy Ministerstvo školství, odpovědnost za získání financí je v současnosti na STK).

---

<sup>3</sup> ŠL – šedá literatura

Oba projekty jsou vzájemně komplementární: zatímco Národní digitální knihovna (jako jádro České digitální knihovny), za jejíž realizaci bude odpovědná Národní knihovna, se orientuje na publikované dokumenty, NUŠL hodlá podchycovat naopak dokumenty oficiálně nepublikované nebo polopublikované.

### **Aktivity odborné komise pro otázky elektronického zpřístupňování VŠKP**

Komise [1] vypracovala a přijala „*Soubor doporučení AKVŠ a Komise eVŠKP pro zpřístupňování VŠKP*“, obsahující základní návody pro implementaci novely VŠ zákona (§47b), workflow, modely rámcových předpisů, metodické návody a metadata [9].

Vysokoškolské kvalifikační práce představují významný, avšak stále jenom dílčí segment obsahu pojmu šedá literatura. Dokumentový záběr NUŠL bude mnohem širší. Úspěchy komise pro VŠKP při přípravě a realizaci uvažovaného Národního archivu VŠKP jsou žádoucím stavem, neboť lze předpokládat, že NUŠL naváže na aktivity vysokých škol v oblasti registrace e-VŠKP. NUŠL bude sklízet metadata a samo bude svá metadata nabízet ke sklizení, přičemž míra spolupráce v segmentu dokumentů e-VŠKP bude záviset na jednotlivých vysokých školách, tedy na stavu jejich digitalizovaných i klasických sbírek a technologické vyspělosti. Klíčovým faktorem zde bude ochota vysokých škol spolupracovat s NUŠL a dále vzájemná technologická kompatibilita jejich dílčích registrů. NUŠL bude koncipováno tak, aby mohlo reflektovat případný další vývoj iniciovaný vysokými školami v této oblasti.

Bez ohledu na případnou realizaci Národního archivu e-VŠKP jsou aktivity komise e-VŠKP a STK komplementární. Normativní dokumenty komise, zejména ty, které se vztahují k formátům, metadatům apod., budou v NUŠL akceptovány a implementovány.

### **Databáze RIV**

Původně uvažovaná možnost přebírat do NUŠL data přímo z Registru informací o výsledcích (RIV) [10] se ukázala jako nerealizovatelná: z databáze RIV lze teoreticky extrahovat data o některých druzích ŠL. Nevýhodou zůstává, že bibliografický popis dokumentů v databázi RIV je nedostatečný. Chybí lokační údaje, bez kterých nelze realizovat základní funkce úložiště. Elektronická ŠL je zachycována pouze okrajově. Věcný popis dokumentů je různorodý a vzájemně nekompatibilní.

Vzhledem k primárnímu účelu databáze RIV – podchycovat výstupy/výsledky projektů a výzkumných záměrů dotovaných státem – nelze předpokládat v dohledné době nějakou zásadní změnu ve struktuře této

databáze. Z hlediska NUŠL má tedy RIV pouze (v nejlepším případě) „referenční“ hodnotu. Pro NUŠL jsou potenciálně daleko zajímavější databáze institucí, které do RIV přispívají, které mohou obsahovat více údajů, než které jsou do RIV odesílány.

Na druhé straně je záměrem tvůrců NUŠL realizace modulu umožňujícího svým účastníkům snadnější spolupráci s RIV. Tento modul by byl schopný generovat z dat o dokumentech ŠL v NUŠL obsažených záznam ve struktuře požadované RIV (týkalo by se pouze těch typů dokumentů, které RIV vyžaduje). Takový modul by byl současně schopen v předstihu reagovat na změny ve struktuře popisu dat pro RIV, která se mění prakticky každý rok.

### ***Základní funkce Národního úložiště šedé literatury***

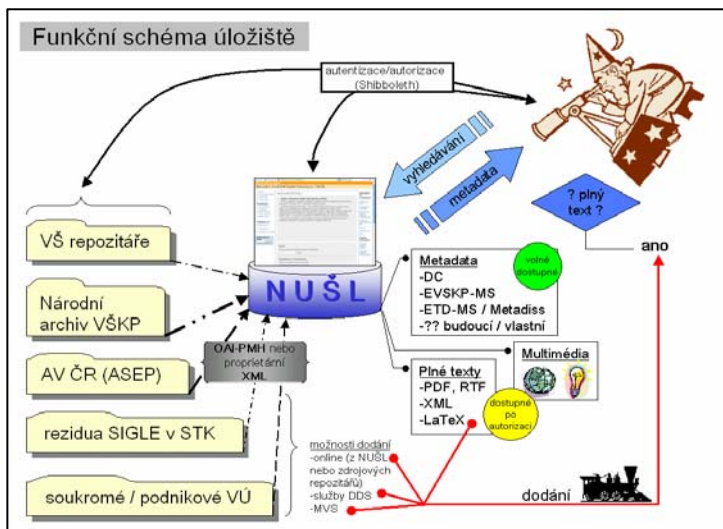
Systém NUŠL bude umožňovat online vyhledání dat o ŠL – **funkce rozhraní** -, existující a podchycené jednak v STK samotné (dílečím úložiště STK), jednak v (dílečích) úložištích zřízených a provozovaných spolupracujícími institucemi, zejména vysokými školami a Knihovnou Akademie věd.

NUŠL bude sklízet data o ŠL z fondů a repozitářů jednotlivých institucí, z takto získaných dat vytvoří databázi metadat. Sklizená metadata budou jak identifikačního a popisného charakteru (**funkce bibliografická**), tak lokačního charakteru, což umožní uživateli zjistit existenci konkrétního tuzemského dokumentu ŠL na území ČR (**lokační funkce souborného katalogu**).

**Funkce rešeršní** bude realizována díky vyhledávacímu rozhraní využívajícímu fulltextový modul indexující plné texty obsažené v úložišti bez ohledu na právo jejich zobrazení či dodání uživateli.

V případě pozitivní identifikace konkrétního dokumentu ŠL umožní uživateli objednat tento dokument u instituce/repozitáře, která jej má v držení (resp. má právo jej zpřístupnit). Pro začátek se předpokládá především sběr metadat, plné texty budou uloženy zejména v lokálních repozitářích, případně také (zrcadlena, popř. archivována) v NUŠL (**depozitní funkce důvěryhodného úložiště**). V případě tradičních (tištěných) dokumentů ŠL bude jako lokační informace sloužit identifikace knihovny ve spojení se signaturou dokumentu.

NUŠL (ani STK) primárně nepřepokládá (ani nevylučuje) retrospektivní digitalizaci dokumentů ŠL. Nedigitalizované dokumenty ŠL budou v NUŠL zachyceny za předpokladu, že jejich metadata budou obsažena v dílečích repozitářích a katalogích spolupracujících institucí a tyto zdroje bude možné obvyklým způsobem sklízet.



Obr.3 – Funkční schéma úložiště

Způsob zpřístupnění dokumentu ŠL bude principiálně závislý na jeho fyzické formě, na omezeních/povoleních daných autorským zákonem a dále na omezeních/povoleních stanovených vlastníkem dokumentu. Samotná transakce zpřístupnění proběhne výhradně mimo NUŠL (tedy mezi uživatelem a dílčím repozitářem/institucí) s výjimkou těch dokumentů, k jejichž zpřístupnění bude oprávněna přímo STK. Případný přístup k plným textům z NUŠL bude omezen podle aktuálních přístupových práv, garantovaných vlastníkem dokumentu (**funkce autorizovaného a selektivního zpřístupnění plných textů**).

Pro zpřístupnění plných textů elektronické ŠL budou podporovány standardní formáty včetně permanentních identifikátorů a definice metadat pro jednotlivé typy ŠL.

Vzhledem ke zkušenostem STK s provozem systémů elektronického dodávání dokumentů lze předpokládat využití také v NUŠL, např. pro definované podмноžiny plných textů, jejichž autorská práva podobný způsob šíření povolují či k jejichž využití dostane STK od držitele autorských práv oprávnění jakéhokoli typu – smluvně či jiným způsobem (**funkce DDS**).

Předpokládanou funkcí bude také možnost sklízet metadata z NUŠL pomocí standardních protokolů, kterou budou moci využít instituce z tuzemska i zahraničí (**funkce otevřeného repozitáře**). Tato služba bude

registrována v příslušných registrech (Registered Data Providers [11], OAster [12], Souborný katalog NDLTD v systému VTLS [13] aj.).

### ***Typy dokumentů podchycovaných NUŠL***

Obsahem úložiště NUŠL bude široká škála dokumentů, které lze (více-méně) řadit do okruhu šedé literatury.

Jde o:

- grantové a výzkumné zprávy (reporty) včetně výstupů nespecifického výzkumu na VŠ,
- VŠKP všech úrovní,
- konferenční materiály a sborníky z konferencí,
- sborníky odborných textů,
- nepublikované učební texty,
- výroční zprávy a jiné interní dokumenty organizací veřejného charakteru,
- programy,
- autorské preprinty a postprinty,
- webovské stránky/blogy odborného charakteru (výběrově).

Uvedené materiály výhradně vědeckého a odborného charakteru budou akceptovány v tradiční tištěné i elektronické podobě. Nepředpokládá se, že doposud (oficiálně) publikované monografie, časopisecké články a stati ze sborníků, které jsou vyhledatelné/ověřitelné v jiných zdrojích, by byly podchycovány.

STK vytipuje z repertoáru dokumentů spadajících pod pojem „publikační činnost“ druhy ŠL, které hodlá prostřednictvím NUŠL podchytit. Lze předpokládat, že jako kritérium bude vytvořena podmínka přijatelnosti vydaných dokumentů do počtu 50 výtisků v případě tištěných dokumentů.

### ***Systémová východiska***

Pro realizaci NUŠL budou vybírány nástroje, koncepce, řešení a prostředky, které se v tomto nebo obdobném kontextu osvědčily a bude možné je po případné úpravě využít.

### ***Programové prostředí a technické řešení***

Přímým vzorem a inspirací pro výběr zde jsou systémy s otevřeným zdrojovým kódem využívané pro provoz institucionálních a oborových repozitářů, např. EPrints, DSpace, Greenstone či Fedora [14, 15, 16, 17]. U těchto systémů je zajištěna podpora pro vysoce heterogenní kolekce metadat, což je klíčové právě pro nasazení v systému uchovávajícím šedou literaturu.

## Komunikační protokoly

Podstatou úložiště bude možnost sklízet metadata (*harvestovat*) a nechat se sklízet (*být harvestován*) pomocí protokolu OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting) [18]. Za účelem kooperace se stávající národní infrastrukturou pro vyhledávání informací v podobě Jednotné informační brány lze předpokládat nasazení komunikačního protokolu kompatibilního se systémem MetaLib, na kterém je Jednotná informační brána založena. Zde se naskýtá možnost využití klasického protokolu Z39.50 [19] či jeho modernějších alternativ založených na komunikaci prostřednictvím XML a běžného protokolu http, jako jsou webové služby (WebServices).

Kompatibilita se stávajícími linkservery (SFX), které jsou v našem prostředí již zavedeny, bude zajištěna odkazováním pomocí protokolu Z39.88-2004 [20] (OpenURL 1.0), který standardně obsahuje i podporu pro protokol OAI.

## Metadatové formáty a schémata

V knihovnách běžně využívané formáty na bázi MARC (MARC 21, UNIMARC) se jeví pro nasazení v případě repozitáře jako nevhodné kvůli své sémantické a syntaktické složitosti a nepřehlednosti. Funkcí repozitáře bude patrně možnost exportu nebo dávkové konverze do těchto formátů v těch případech, kde úroveň popisu bude odpovídat cílovému formátu.

Primárním metadatovým formátem však bude běžně nasazovaný a využívaný formát nekvalifikovaného Dublinského jádra (Dublin Core) [21], které bude se svými 15 jednoduchými elementy tvořit nejnižší společný jmenovatel pro všechny typy dokumentů ŠL obsažené v úložišti.

Pro specifické skupiny dokumentů lze předpokládat implementaci sémanticky propracovanějších formátů, které respektují odlišnosti jednotlivých typů dokumentů. Pro VŠKP lze očekávat podporu metadatového schématu připravovaného Komisí pro vysokoškolské kvalifikační práce (EVSKP-MS) [22]. Akceptovatelné však budou i další formáty pro VŠKP (ETD-MS či Metadiss), neboť nelze vyloučit sklizení metadat dokumentů českých autorů z úložišť mimo ČR (EU, USA, ...). Pro další typy ŠL lze předpokládat nasazení dalších (budoucích, existujících či vlastních) metadatových formátů. Jako nejvhodnější se jeví aplikace dle metadatového schématu MODS [23] či vyvinutí vlastní verze tohoto schématu, která by respektovala specifika šedé literatury.



## **Permanentní identifikátory**

Jako nutnost se jeví zavedení permanentních (persistentních) identifikátorů, sloužících k jednoznačnému odkazování na plné texty. V úvahu připadají zejména identifikátory Handle/DOI [24, 25].

## **Možnosti bibliografického popisu**

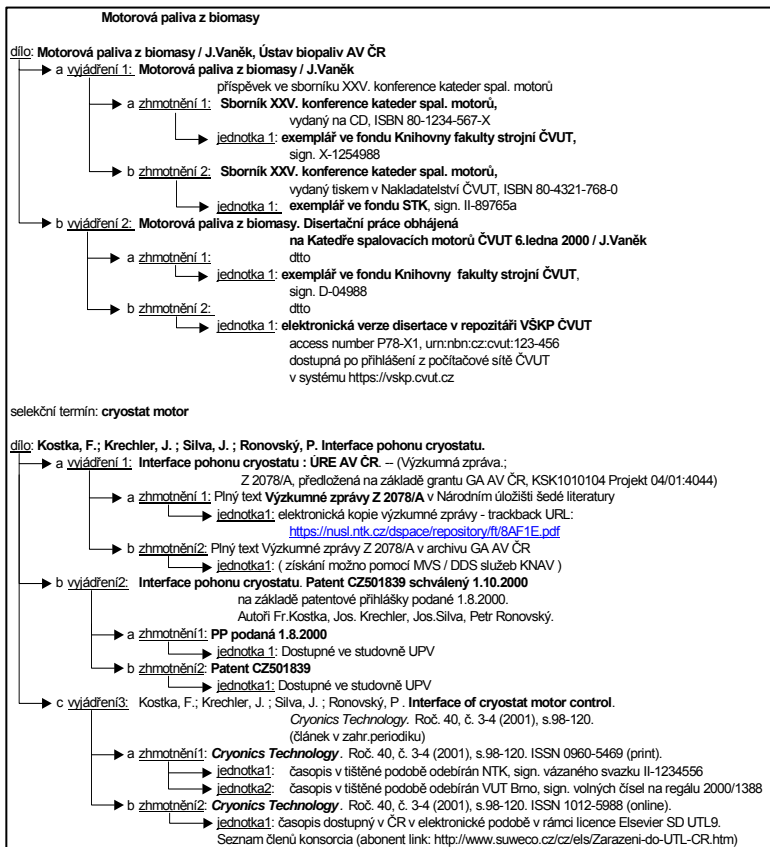
Oblast šedé literatury je díky své podstatě více než vhodná pro nasazení moderních formátů pro bibliografický popis, realizovaných Programem UBCIM (IFLA). Zde se přímo naskýtá vhodnost nasazení entit, atributů a vazeb známých jako Funkční požadavky pro bibliografické záznamy (FRBR) [26]. Použití FRBR by patrně výrazně zvýšilo využitelnost NUŠL. Více o možnostech FRBR vypovídá tabulka č.1 s fiktivními daty<sup>4</sup>.

## **Věcný popis**

Otázka věcného popisu zůstává stále otevřena, jednou z možností je využití Seznamu předmětových kategorií SIGLE (SIGLE Subject Category List) doplněného o předmětová hesla Polytematického strukturovaného hesláře (PSH) a volně tvořená autorská klíčová slova. Druhá možnost spočívá v ponechání prvků předmětové klasifikace tak, jak byla převzata ve sklizených záznamech.

---

<sup>4</sup> Ukázka zahrnuje dokumenty, které nejsou zařazovány do sféry šedé literatury (články z publikovaných časopisů, či dokumenty specifického charakteru, které do sféry šedé literatury zařadit lze (patenty, patentové přihlášky), které však NUŠL dle stávající verze konceptu zahrnovat nebude.



Tabulka 1 – ukázka možného uspořádání výsledků vyhledávání s využitím FRBR

## Autentikace a autorizace<sup>5</sup>

Bezespору klíčovou otázkou, na které závisí úspěch NUŠL, je otázka zabezpečení oprávněnosti přístupu. STK bude sledovat existující aktivity budování jednotné architektury pro ověřování identity, na které kooperuje sdružení CESNET ve spolupráci s vysokými školami. Lze předpokládat, že NUŠL bude jednou z důležitých aplikací, využívajících technologii Shibboleth [27].

<sup>5</sup> autentikace – zjištění kdo je osobou požadující službu (=identifikace)  
 autorizace – přidělení oprávnění dané osobě požadující službu

## **Formáty vkládání, ukládání a zpřístupňování**

Formáty plných textů bude nutno rozlišit dle kontextu workflow na formáty vstupní, úložné a výstupní.

Na vstupu úložiště budou rozlišovány podporované – HTML, plaintext ASCII, DjVu, PS, XML, LaTeX, PDF, RTF, Odx (OpenOffice) a akceptované bez záruky (DOC).

Dokumenty budou ukládány v XML kontejnerech dle standardu METS [28], což je v současnosti nejpracovanější standard formátu pro digitální úložiště.

Výstupním formátem bude jak verze v originálním formátu, tak pomocí XSLT stylů na vyžádání generovaná verze ve formátu PDF.

## **Funkcionalita důvěryhodného úložiště**

Lze předpokládat, že autoři a instituce, vkládající svá díla do NUŠL, budou požadovat maximální kvalitu služeb úložiště. Jako optimální se jeví využít existujících standardů, definujících služby tzv. důvěryhodného úložiště (trusted repository). Tyto standardy a navazující atributy byly stanoveny v letech 2000-2002 společnou aktivitou OCLC a RLG (Research Libraries Group). Vypracovaná koncepce se stala referenčním modelem OAIS (Open Archival Information System Reference Model) [29] a posléze standardizována jako norma ISO. Norma definuje vlastnosti úložiště a jednotlivé odpovědnosti zahrnující administrativní, finanční, organizační, technologické, bezpečnostní a procedurální postupy. Jejich respektování umožňuje úložiště certifikovat jako důvěryhodné. Díky náročnosti splnění jednotlivých odpovědností je možné, že vlastnosti důvěryhodného úložiště budou implementovány postupně, po částech.

## ***Zajištění projektu NUŠL***

### **Řešitelská kapacita**

Pro řešení tohoto projektu vytvořila STK řešitelský kolektiv cca 5-6 pracovníků se souhrnnou kapacitou cca 1,3 úvazku./rok.

### **Finanční zajištění**

Náklady na řešení projektu v roce 2006 jsou hrazeny z rozpočtu STK. Výraznější investice zejména do programového vybavení/technologie se předpokládají v roce 2007 a v následujících letech. STK bude usilovat primárně o jejich pokrytí z prostředků svého rozpočtu, v případě nedostatečného pokrytí i z jiných zdrojů (resortní výzkum).

## Spolupráce

Během prací na projektu NUŠL předpokládá STK spolupráci s následujícími institucemi:

- Knihovna AV ČR (KNAV),
- Komise pro e-VŠKP,
- vysoké školy provozující či vyvíjející vlastní repozitáře,
- ostatní vysoké školy (potenciální zdroje metadat),
- Ústav informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze (ÚISK FF UK),
- Národní knihovna České republiky (NK ČR).

## Závěr

Je evidentní, že informační potenciál šedé literatury je v ČR opomíjen. Přípravované projekty národních centralizovaných archivů a úložišť mají za cíl tento stav změnit. Lze předpokládat, že poptávka po získání dokumentů šedé literatury poroste a navrhovaný způsob v podobě Národního úložiště šedé literatury je nejlepším prostředkem, jak tuto poptávku uspokojit. Státní technická knihovna (a její budoucí nástupce Národní technická knihovna) se bude snažit nejen o navržení a realizaci celého projektu, ale i o kontinuální zabezpečení a údržbu běžného provozu systému.

*Za pomoc při zpracování tohoto příspěvku děkuji PhDr. Janu Bayerovi, Mgr. Lence Laňkové, Bc. Lindě Skolkové a Ing. Martinu Svobodovi.*

## Literatura:

[1] Odborná komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací. <http://www.evskp.cz>

[2] DE ROSA, Cathy, DEMPSEY, Lorcan, WILSON, Alane. *The 2003 OCLC environmental scan : pattern recognition : executive summary*. Dublin, Ohio : OCLC, 2004. 17 s. Dostupný z WWW: <http://www.oclc.org/reports/escan/>. ISBN 1-55653-355-1.

[3] Komise Evropských společenství. *I2010: Digitální knihovny : sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů*. Brusel : KOMISE EVROPSKÝCH SPOLEČENSTVÍ, 2005. 12 s. Dostupný z WWW: [http://europa.eu.int/information\\_society/activities/digital\\_libraries/doc/communication/cs\\_comm\\_digital\\_libraries.pdf](http://europa.eu.int/information_society/activities/digital_libraries/doc/communication/cs_comm_digital_libraries.pdf).

[4] DEWATRIPONT, Mathias, et al. *Study on the economic and technical evolution of the scientific publication markets in Europe : final report - January 2006*. Brussels : European Commission, 2006. 112 s. Dostupný z

- WWW: [http://ec.europa.eu/research/science-society/pdf/scientific-publication-study\\_en.pdf](http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf) ISBN 92-79-01029-8.
- [5] *Dlouhodobý záměr vzdělávací a vědecké, výzkumné, vývojové, umělecké a další tvůrčí činnosti pro oblast vysokých škol na období 2006 – 2010*. Praha : Ministerstvo školství, mládeže a tělovýchovy, 2005. 39 s. Dostupný z WWW: [http://www.msmt.cz/Files/PDF/DZ\\_SWOT\\_30\\_8\\_05.pdf](http://www.msmt.cz/Files/PDF/DZ_SWOT_30_8_05.pdf).
- [6] Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010. *Knihovna plus* [online]. 2006, č. 1. Dostupný z WWW: [http://knihovna.nkp.cz/knihovna61/koncepce\\_uchovavani.htm](http://knihovna.nkp.cz/knihovna61/koncepce_uchovavani.htm). ISSN 1801-5948.
- [7] Česko. Zákon ze dne 20. prosince 2005, kterým se mění zákon č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a některé další zákony. In Sb. zákonů. 2005, částka 188. Dostupný také z WWW: [http://www.msmt.cz/Files/vysokeskoly/Legislativa/Novela\\_zakona\\_552\\_20\\_05.htm](http://www.msmt.cz/Files/vysokeskoly/Legislativa/Novela_zakona_552_20_05.htm).
- [8] Česko. Zákon ze dne 25. dubna 2006, kterým se mění zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, a některé další zákony. In Sb. zákonů. 2006, částka 72. Dostupný také z WWW: [http://www.nkp.cz/o\\_knihovnach/AutZak/Index.htm](http://www.nkp.cz/o_knihovnach/AutZak/Index.htm).
- [9] Soubor doporučení AKVŠ a Komise eVŠKP pro zpřístupňování VŠKP. viz <http://www.evskp.cz/sd.php>
- [10] Databáze „Registr informací o výsledcích státem podporovaného výzkumu a vývoje“ – RIV. viz <http://www.vlada.cz/rvv/riv/index.html>
- [11] Registered Data Providers. viz <http://www.openarchives.org/Register/BrowseSites>
- [12] OA Ister. viz <http://oaister.umdl.umich.edu/o/oaister/>
- [13] Networked Digital Library of Theses and Dissertations Union Catalog. viz <http://hercules.vtls.com/cgi-bin/ndltd/chameleon>
- [14] Greenstone Digital Library Software. viz <http://www.greenstone.org/>
- [15] Dspace Federation. viz <http://www.dspace.org/>
- [16] Eprints.org. viz <http://www.eprints.org/>
- [17] FEDORA Digital Repository System. viz <http://www.fedora.info/>
- [18] OAI. Open Archives Initiative. viz <http://www.openarchives.org/>
- [19] Z39.50 Implementors Group - Czechia. viz <http://www.stk.cz/ZIG>
- [20] Z39.88-2004. viz [http://www.niso.org/standards/standard\\_detail.cfm?std\\_id=783](http://www.niso.org/standards/standard_detail.cfm?std_id=783)
- [21] Dublin Core Metadata Initiative. viz <http://dublincore.org/>

- [22] EVSKP-MS - Metadatový soubor pro elektronické vysokoškolské kvalifikační práce v ČR. viz <http://www.evskp.cz/standardy/metadata/>
- [23] MODS. Metadata Object Description Schema. viz <http://www.loc.gov/standards/mods/>
- [24] The Handle System. viz <http://www.handle.net/>
- [25] DOI. The Digital Object Identifier System. viz <http://www.doi.org/>
- [26] IFLA Study Group on the Functional Requirements for Bibliographic Records. Funkční požadavky na bibliografické záznamy : závěrečná zpráva. Ludmila Celbová. München : Saur, 1998. 69 s. UBCIM publications. New Series; sv. 19. Dostupné z <http://webarchiv.nkp.cz/frbr.pdf>.
- [27] Shibboleth. viz <http://shibboleth.internet2.edu/>
- [28] METS. Metadata Encoding & Transmission Standard. viz <http://www.loc.gov/standards/mets/>
- [29] OAIS. Reference Model for an Open Archival Information System. viz <http://public.ccsds.org/publications/archive/650x0b1.pdf>