

IDENTIFIKACE ZDROJŮ A VYHLEDÁVÁNÍ INFORMACÍ

SOUHRN

TÉMATICKÉ VYMEZENÍ

Tato kapitola se zabývá otázkami popisu zdrojů v souvislosti s narůstajícím počtem dokumentů v elektronické podobě včetně metadat, řízených slovníků a tezurů, vyhledávacích nástrojů, bran, XML a Z39.50.

POLITICKÉ SOUVISLOSTI

Rozvoj Internetu znamená, že knihovny mají nyní přístup ke stále většímu objemu informací v elektronické podobě. Pracovníci knihoven by tedy měli vědět, jak svým uživatelům co nejlépe poradit v oblasti identifikace a vyhledávání potřebných informací v digitálním prostředí. Informační společnost, celoživotní vzdělávání a zvyšující se tendence k propojení se státními orgány na elektronické bázi - všechny tyto okolnosti zvyšují důležitost vyhledávání informací pro všechny občany. Klíčovou otázkou je zde univerzálnost objemu digitálních informací.

UŽITEČNÉ INFORMACE PRO ÚSPĚŠNOU PRAXI

Tato část kapitoly je rozdělena do tří bodů:

Popis zdrojů: jsou uvedeny různé prostředky popisu zdrojů za účelem zvýšení přesnosti vyhledávání informací, např. metadata či systémy řízených slovníků jako jsou tezury a speciální identifikační slovníky.

Identifikace a vyhledávání zdrojů: jsou uvedeny prostředky, kterými se vyhledávají a identifikují zdroje, např. vyhledávací nástroje, informační brány, skupiny, rejstříky veřejných informací. Identifikace zdrojů je logicky druhotná. Byl-li předmět metodicky popsán, nástroj, který ho metodicky vyhledává má větší šanci jej nalézt. Identifikace a vyhledávání zdrojů je oblast, ve které mají knihovny příležitost ke spolupráci s ostatními organizacemi, například středními a vysokými školami a mohou tak poskytovat lepší přístup ke zdrojům. Partneři se obvykle dohodnou na společném poskytování služeb uživatelům a recipročním přístupu do katalogů.

Technologie s tím spojené: jsou uvedeny například XML, RDF, Z39.50, vyhledávání obrázků včetně CBIR, ontologie-formální popis objektů a jejich vzájemných vztahů, tematické mapy, které zachycují určitou oblast.

ÚKOLY PRO BUDOUCNOST

V budoucnu budou lidé obklopeni inteligentními, komunikativními a spolehlivými přístroji, které budou schopny reagovat jako lidské bytosti. Technologie popsané v této kapitole již byly vyvinuty, ale jejich vzájemná interakce je stále otázkou budoucnosti. Ovlivní domácnosti, školy, hotely, auta, letadla - zkrátka všechny aspekty života. Jejich vliv na knihovny a jiné kulturní organizace bude dalekosáhlý.

TÉMATICKÉ VYMEZENÍ

Otázky, kterými se tato kapitola zabývá, jsou:

Popis zdrojů

Metadata

Popisy na úrovni sbírek

Řízené slovníky a tezury

Speciální identifikační slovníky

Vyhledávání zdrojů

Knihovní katalogy

Nástroje vyhledávání

Seznamy doporučených stránek

Informační brány a portály

Sdružení

Místní informace

Podpora vzdělávání

XML

RDF

Z39.50

Harvest
Vyhledávání obrázků
Semantický web
Ontologie
Digitální podpisy
Webové služby
Odkazy na nejlepší praktické zkušenosti

Poznámka: Očekává se, že knihovny zaujmou postoj k popisu zdrojů, jejich identifikaci a vyhledávání z různých výchozích pozic a mnohé z nich, ne-li všechny, budou mít potřebu přijmout a modifikovat univerzální, již připravený postup. Následující informace mohou pomoci buď při pořizování knihovnických systémů nebo jednáních s dodavateli či odbornými konzultanty.

POLITICKÉ SOUVISLOSTI

Závažnost vymezené problematiky pro knihovny

Knihovny působí v síťovém informačním prostředí. Rozvoj Internetu v praxi znamená, že dokonce i ta nejmenší pobočka knihovny má přístup ke stále většímu objemu informací poskytovaných v elektronické podobě. Pracovníci knihoven by tedy měli vědět, jak svým uživatelům co nejlépe poradit v oblasti identifikace a vyhledávání potřebných informací v digitálním prostředí. Informační společnost, celoživotní vzdělávání a zvyšující se tendence ke komunikaci se státními orgány na elektronické bázi - všechny tyto okolnosti zvyšují důležitost vyhledávání informací pro všechny občany.

Řada knihoven vytváří také nové digitální dokumenty sama a to buď na svých webových stránkách, či novým multimediálním obsahem. Někdy jsou tyto projekty financovány v rámci specifických programů digitalizace /viz kap. Digitalizace a Financování/. Knihovny by také měly zvážit, jak nejlépe tento digitální obsah popsat, aby byl jak lehce vyhledatelný, tak také kompatibilní s jinými sbírkami digitálních dokumentů .

Technologie a normy jsou v této oblasti stále ve stádiu vývoje a během času se budou měnit a dále rozvíjet. Knihovny musejí v této oblasti sledovat každou novinku, aby se vyhnuly nevyhovujícím či nevhodným technologiím a standardům.

Knihovny musejí těmto otázkám dobře rozumět, aby mohly plánovat práci a stanovit priority, ať už jde o počáteční stádium využívání nových systémů nebo jejich další rozvoj ve spolupráci s externími konzultanty.

Co je univerzálnost?

V souvislosti se soubory digitálních dokumentů znamená univerzálnost možnost širšího opakovaného využití, a přenosnosti určitých částí nebo zařízení, například v rámci různých sítí, systémů a organizací po dobu co nejdéle. Klíčem k dosažení toho cíle jsou standardy - kodifikovaná pravidla a zásady tvorby, popisu a správy digitálních zdrojů /viz Reinventing the Wheel – D-Lib magazine, leden 2002 <http://www.dlib.org/dlib/january02/gill/>.

Proč je univerzálnost důležitá?

Aby uživatelé mohli snáze vyhledávat a získávat informace z různých zdrojů - katalogů, z různých institucí /knihoven, muzeí a archivů/ a také z různých typů zdrojů /knih a muzejních artefaktů/.

Nové služby a různé druhy spolupráce

Zpřístupňování zdrojů prostřednictvím Internetu, ať už jsou vytvořeny samotnou knihovnou, nebo získány jinou cestou /spolupráce, koupě, licence apod./, nepřímě naznačuje jaké pracovníky knihovna potřebuje; jakým školením by měli projít; jaké služby bude poskytovat; a jak budou tyto služby poskytovány uživatelům /viz [multimedia](#)/. Knihovny budou muset na místní úrovni učinit strategická rozhodnutí - stanovit priority, výši investic a cíle.

Zvyšující se závislost na digitálním obsahu

Knihovny budou zpracovávat stále větší množství dokumentů nezbytně důležitých pro uživatele, které jsou přístupné pouze v elektronické podobě. Tyto materiály mohou být dražší, než jejich tištěný ekvivalent a podléhat vydavatelské či distributorské kontrole a dozoru. Důležitou roli zde bude hrát také licencování.

DOPORUČENÍ PRO ÚSPĚŠNOU PRAXI- Popis zdrojů

Knihovny by měly vzít v úvahu aplikaci metadat do zdrojů a webových stránek, které vytvářejí. Měly by také zvážit využití XML.

Metadata

Stručně mohou být definována jako "strukturovaná data o datech". Například zápis v katalogu je metadatem, které popisuje určitou knihu. Metadata se začala používat jako prostředek zdokonalení efektivnosti vyhledávání digitálních zdrojů na síti, a to zavedením ucelené struktury pro popis webových stránek a dalších digitálních zdrojů. Metadata spojená s popisem knih mohou zahrnovat tyto prvky: autora, titul, vydavatele, datum, ISBN, klasifikaci atd; metadata spojená s popisem webových stránek pak : titul, tvůrce, předmět (předmětové heslo) popis atd. Pro úvod do problematiky je velmi užitečná webová stránka Diffuse <http://www.diffuse.org/meta.html>.

Metadata mohou být využita k různým účelům:

- **Metadata popisu zdrojů** – slouží k popisu zdrojů a jejich vyhledávání. Je k dispozici celá řada různých druhů, ze kterých jsou patrně nejznámější Marc a Dublin Core:

Marc - strojem čitelná katalogizace /Machine Readable Cataloguing /<http://www.locweb.loc.gov/marc/>, je bibliografické metadatové schéma provozované Kongresovou knihovnou ve Washingtonu D.C. Současnou verzí je MARC 21.

Dublin Core <http://dublincore.org/> - systém popisu "dokumentů jako předmětů", který se skládá z následujících patnácti metadatových elementů: 1. titul 2.tvůrce 3. předmět 4. předmětová hesla, klíčová slova 5. vydavatel 6. přispěvatel 7. datum 8. typ zdroje 9. formát 10. identifikátor zdroje 11. zdroj 12. jazyk 13. příbuzné vztahy 14. pokrytí 15. management práv k dílu.

- **Metadata pro oblast ochrany dokumentů** - slouží k podpoře ochrany dokumentů a aktivit spojených s činnostmi v oblasti archivnictví. Projekt CEDARS <http://www.leeds.ac.uk/cedars/> se zabývá otázkami spojenými s ochranou digitálních zdrojů.
- **Metadata pro oblast autorských práv** – slouží k popisu otázek souvisejících se zákonem o duševním vlastnictví /IPR/ ke zdroji. Například viz projekt Indecs <http://www.indecs.org/>.

- **Metadata pro oblast vzdělávání** - mají napomoci s úkoly vzdělávacích institucí v oblasti vyhledávání zdrojů a správou agend virtuálního vzdělávacího prostředí (např. vedení záznamů o studentech a popis kurzů). Pro více informací o metadatach podporujících elektronické vzdělávání si můžete najít webovou stránku Diffuse <http://www.diffuse.org/eLearning.html>.

- **Zeměpisná metadata** - používají se při popisu digitálních map a zeměpisných informačních systémů /Viz webová stránka Diffuse <http://www.diffuse.org/gis.html#help> - informace o normách pro výměnu zeměpisných dat/.

Specifická oborová metadata byla vytvořena na základě specifických požadavků v určitých oblastech, například:

- **Archívy** - tam se běžně užívá EAD <http://lcweb.loc.gov/ead> - kódovaný archivářský popis /Encoded Archival Description/, nebo ISAD/G/<http://lcweb.loc.gov/ead> - - obecná mezinárodní norma pro archivářský popis /General International Standard Archival Description/. O normách pro archívy se můžete více dovědět také na webové stránce Diffuse <http://www.diffuse.org/archives.html>.
- **Muzea** – CIMI <http://www.cimi.org> je mezinárodní konsorcium institucí uchovávajících kulturní dědictví, které spolupracují a podporují společně rozvoj a aplikaci příslušných norem pro muzejní informace.O

normách pro muzea se můžete více dovědět také na webové stránce Diffuse <http://www.diffuse.org/museum.html>.

Vláda – GILS <http://www.gils.net/about.html> neboli Vládní informační servis/Globální lokátor informací /Government Information Locator Service/Global Information/ je široce používán pro informace státní správy, i když se v poslední době zdá, že řada vlád začíná preferovat Dublin Core. Iniciativa pro metadata Dublin Core (DCMI – Dublin Core Metadata Initiative) založila Vládní pracovní skupinu (Government Working Group) <http://dublincore.org/groups/government>. Řada vlád, například ve Velké Británii, Austrálii, Kanadě, USA či Dánsku zpracovala směrnice, které mohou být pro organizace veřejného sektoru povinné, což se samozřejmě týká také veřejných knihoven. Viz e-gif /Electronic Government Interoperability Framework <http://www-envoy.gov.uk/publications/frameworks/egif4/egif4.htm/> a e-gms /Electronic Government Metadata Standard <http://www-envoy.gov.uk/publications/guidelines/metadata/index.htm/> / vytvořený pobočkou e-Envoy ve Velké Británii <http://www-envoy.gov.uk/index.htm/> . První verze e-gms byla založena na prostém Dublin Core, druhá se však posunula ke kvalifikovanému Dublin Core s některými přidavnými dokumenty pro účely řízení .

Popisy na úrovni sbírek

V rozsáhlých projektech spolupráce je často nezbytné použít k popisu fondů spolupracujících knihoven metadata souborné úrovně /tematické vymezení, úroveň, hloubka, jazyk atd./ . Viz např. Zaměření na popis sbírek (Collection Description Focus) <http://www.ukoln.ac.uk/cd-focus> v rámci britského UKOLN pro informace o využití popisu na úrovni sbírek v oblasti vyššího a dalšího vzdělávání ve Velké Británii a Směrnice NOF <http://www.ukoln.ac.uk/nof/support>, které by měly přijmout veřejné knihovny.

Řízené slovníky a tezaury

Metadata jsou jen polovičním řešením. Jestliže mají být uživatelé schopni provádět v distribuovaných datech rozsáhlá vyhledávání, pak do nich musí tvůrci vložit elementy metadat v ucelené podobě. Aby to byli schopni provést, musí si pro indexování osvojit a používat nějaké identifikovatelné kódovací schéma či řízený slovník. Dobrým příkladem jsou Předmětová hesla Kongresové knihovny (LCSH – Library of Congress Subject Headings).

Tezaurus je řízený slovník, kde jsou termíny zařazené do hierarchie, které ukazují jejich vzájemné vztahy jako širší či užší termíny, ekvivalenty či částečné ekvivalenty, preferované či nepreferované termíny /pro kontrolu synonym/ . Často také obsahují poznámky k tematické vymezení a jiné důležité informace. Pro tezaury existují dvě normy ISD: *ISO 2788, 1986 Průvodce založením a rozvojem jednojazyčných tezurů; a ISO 5964, 185 Průvodce založením a rozvojem vícejazyčných tezurů.*

Webová stránka Gettyho muzea uveřejňuje řadu tezurů, například Umělecký a architektonický tezaurus <http://www.getty.edu/research/tools/vocabulary/aat/> a Tezaurus zeměpisných názvů <http://www.getty.edu/research/tools/vocabulary/tgn/>. Traugott Koch vytvořila užitečný seznam řízených slovníků, tezurů a klasifikačních schémat <http://www.lub.lu.se/metadata/subject-help.html>.

Jedinečné identifikátory

Aby byl každý zdroj na síti dobře k nalezení, musí obsahovat jedinečný identifikátor nebo adresu, jako například URL /Uniform Resource Locator - univerzální lokátor zdrojů/. Problém s URL, jak každý uživatel Internetu jistě ví, spočívá v tom, že každý udaný zdroj se pravděpodobně časem změní či přemístí, což vyhledávání informací podstatně stěžuje. DOI <http://www.doi.org/> /Digital Object Identifier - identifikátor digitálních objektů/ je jedním z identifikátorů, který tento problém pomáhá překlenout. Paul Miller napsal dobře srozumitelný článek <http://www.ariadne.ac.uk/issue24/metadata/intro.html> objasňující problematiku těchto jedinečných identifikátorů.

DOPORUČENÍ PRO ÚSPĚŠNOU PRAXI- Identifikace a vyhledávání zdrojů

Vyhledavače

Knihovní katalog je klíčem k fondům každé knihovny. Avšak mnoho, ne-li většina knihoven umožňuje svým uživatelům přístup k Internetu. Musí tedy nalézt způsob, jak pomoci zpřístupnit uživatelům tyto síťové zdroje. Pro knihovny se zdá nejjednodušším způsobem prostě upozornit uživatele na nejlepší vyhledavače. K dispozici jsou následující typy:

- **Vyhledavače volného textu**, které hledají na základě klíčových slov, např. Altavista <http://www.altavista.com>, Google <http://www.google.com>
- **Vyhledavače pracující na principu rejstříku a adresáře** ve kterých uživatel zužuje svůj výběr pomocí různých menu, např. Yahoo
- **Meta/multi vyhledavače**, kdy se vyhledává na více než jednom vyhledávači, např. ixquick <http://www.ixquick.com/>.
- **Vyhledavače pracující na principu přirozeného jazyka** - zde může být vyhledávání zadáno pomocí otázky, např. Ask Jeeves <http://www.ask.co.uk/>.

Téměř vyčerpávající seznam všech typů vyhledavačů naleznete a více informací o nich se dovíte na www.philb.com a <http://searchenginewatch.com>. Tyto stránky podávají informace o vyhledávání na Internetu, analyzují produkci a parametry vyhledavačů a pomáhají majitelům webových stránek se zlepšením jejich vyhledatelnosti na síti.

Seznam doporučených stránek

Knihovny mohou uživatelům ještě lépe pomoci vytvořením seznamu užitečných webových stránek. Obvykle jsou hodnoceny podle nějakého kvalitativního kritéria a často členěny na základě obsahu <http://www.desire.org/handbook/2-1.html>. RND - Resource Discovery Network <http://www.rdn.ac.uk/> - síť odhalování zdrojů - je tohoto dobrým příkladem. Kontrola a aktualizace těchto seznamů mohou být však časově náročné. Knihovny by proto měly zvážit při vytváření a aktualizaci těchto seznamů možnost spolupráce. CORC (Cooperative Online Resource Catalog Service) <http://www.oclc.org/corc/> od OCLC je dobrým příkladem spolupráce při vytváření popisů webových stránek. /Viz také portály/.

Informační brány a portály

Rozdíl mezi informačními branami a portály je velice nejasný, ale informační brána se obvykle skládá ze souboru anotovaných odkazů na další webové stránky, které byly jejími tvůrci vybrány. Uživatel může dostat k dispozici celou řadu menu, které jej dovedou od obecných ke specifitějším pojmům. Takovéto stránky jsou pro výzkumné pracovníky nesmírně užitečné, protože jim usnadňují práci při hledání jedné webové stránky s určitým tématem za druhou, ale umožňují jim vybrat právě ty nejkvalitnější. Portály mají obecně za úkol poskytovat uživatelům kromě souboru odkazů také ostatní služby. Poskytují příležitosti k on-line transakcím různého druhu jako je například on-line nákup atd. Mohou podporovat také prohledávání dokumentů na stránkách, které spojuje jednotná úprava.

Není možné uvést seznam všech informačních bran, zde jsou tedy alespoň příklady některých z nich:

- **Obecné** www.yahoo.com - série menu pokrývajících World Wide Web jako celek
- **Vládní** www.ukonline.gov.uk tématicky a abecedně seřazený seznam webových stránek vlády Velké Británie a místních samospráv, www.premier-ministre.gouv.fr vstupní brána do vládních úřadů francouzské vlády.
- **Akademické** www.rdn.ac.uk - síť identifikace zdrojů (RND- Resource Discovery Network) vznikla na podporu potřeb vyššího a dalšího vzdělávání ve Velké Británii. RND je fakticky síť samostatných vstupních bran, přičemž každá z nich pokrývá určitou tématickou oblast.
- **Knihovnické** portál knihoven UNESCO http://www.unesco.org/webworld/portal_bib/
- **Muzejní** Nonstop muzeum <http://www.24hourmuseum.org.uk/>
- **Archivní** portál archivů UNESCO http://www.unesco.org/webworld/portal_archives/pages/index.shtml
- **Kulturní** nový **Evropský kulturní portál** při **Evropské komisi** http://europa.eu.int/comm/culture/index_en.htm
- **Ostatní** www.teldir.com - seznam odkazů na on-line telefonní seznamy.

K podpoře rozvoje a udržování informačních bran je k dispozici software, například ROADS <http://www.roads.int.ac.uk/index>. (Resource Organisation and Discovery in Subject based services – Organizace a vyhledávání zdrojů v předmětově založených službách). Komerční společnosti také nabízejí portálový software, např. Vignette <http://www.vignette.com>, SAP Portal <http://www.sapportals.com/>, Plumtree <http://plumtree.com>

atd. Některé knihovny začínají vytvářet své vlastní informační brány a portály - například k podpoře tzv. "sdružení", místních informačních služeb a poskytování podpory vzdělávání.

Clumps (sdružení, konsorcia)

"Sdružení" je skupina institucí, obvykle veřejných knihoven, akademických knihoven a jiných organizací /například muzeí a archivů/, které se dohodnou na spolupráci v určitých oblastech a vytvoří např. konsorcium pro nákup zdrojů /zejména elektronických zdrojů/, meziknihovní výpůjční službu nebo vzájemné propojení katalogů. Důležitým rysem většiny sdružení je, že vytvářejí informační bránu či portál, který umožňuje uživatelům vyhledávání v databázích všech propojených knihoven současně. Aby to bylo možné, musí se sdružení dohodnout na společném metadatovém profilu, podle kterého může být používání různých struktur metadat v rámci konsorcia mapováno. Řada profilů byla vyvinuta speciálně pro podporu těchto aktivit, například "Bath Profile" nebo "ONE-2 Profile". Z39.50 nebo XML se používají k vyhledávání mezi různými katalogy a/nebo databázemi. Licencování a také záležitosti týkající se autorských práv různých elektronických zdrojů přístupných prostřednictvím sdružení mohou být problematické.

Místní informace

Rada hrabství v Essexu ve Velké Británii je průkopníkem rozvoje distribuovaného systému místních informací. Místo tradiční ústřední databáze místních informací Essex Online <http://www.seamless.org.uk/> která pochází z projektu SEAMLESS/ používá k vyhledávání databázi a webových stránek spolupracujících organizací protokolu Z39.50 a Harvest v jednom integrovaném vyhledávání, což je také známo jako hluboká integrace. Přes 30 místních organizací /městské rady, univerzity, obchodní agentury, zdravotnické organizace, místní novinářská společnost, dobrovolné agentury atd./ v současné době zpřístupňuje svá data pro tento systém a jejich počet den ode dne roste.

Aby systém mohl v těchto distribuovaných souborech dat vyhledávat, dohodli se spolupracující organizace na společném profilu metadat založených na e-gms http://www.govtalk.gov.uk/interoperability/metadata_document.asp?docnum=524, (který vychází ze souboru Dublin Core) a na základním tezauru pro popis. Rada hrabství Essex nedávno získala grant od New Opportunities Fund /Fondu nových příležitostí/, na rozvoj systému v národním měřítku, zavedení dat od velkých národních poskytovatelů informací jako je např. nhsDirect Online <http://www.nhsdirect.nhs.uk/> a zpřístupnění systémových prostředků pro vyhledávání informací v dalších územních celcích spravovaných 8 místními úřady, což představuje služby asi pro 6 milionů lidí. Více informací viz webová stránka [seamlessUK](http://www.seamlessUK).

Podpora vzdělávání

Některé veřejné knihovny začínají spolupracovat s jinými organizacemi, vysokými školami atd., aby byly schopny poskytnout lepší podporu studentům. Spolupracující organizace se obvykle dohodnou na vzájemném zpřístupnění svých knihoven /a také služeb/ uživatelům a mohou zavést jednotný knihovní průkaz. Často vyvíjejí systémy umožňující jednotné vyhledávání ve všech jejich katalozích a/nebo databázích /viz výše zmíněné sdružení/. Příklady zahrnují: Libraries Access Sunderland scHeme <http://www.lash.sunderland.ac.uk/> a Glasgowský Digital Library Project <http://gdlr.strath.ac.uk/>.

Sdružení výkonných rad informačních služeb (JISC) ve Velké Británii vyvíjí Distribuovaný národní vzdělávací zdroj (DNER) k poskytování informací a zdrojů k podpoře výzkumu a vzdělávání. To je velmi rozsáhlý a revoluční projekt, který se zakládá na myšlence tematických portálů a je průkopníkem v oblasti zajišťování zdrojů v síťovém informačním prostředí.

Řada akademických knihoven je zapojena také do rozvoje Řízeného nebo Virtuálního učebního prostředí /MLE, VLE/. Ve Velké Británii vede UKOLN Skupinu pro metadata ve vzdělávání /MEG – Metadata for Education Group <http://www.ukoln.ac.uk/metadata/education/>, aby zvážil důsledky zavedení metadat na podporu těchto vzdělávacích systémů. Veřejné knihovny mohou v tomto procesu hrát také důležitou úlohu, zejména s ohledem na jejich roli v podpoře formálního a celoživotního vzdělávání.

DOPORUČENÍ PRO ÚSPĚŠNOU PRAXI

Základní technologie

XML

Extenzivní Mark-up jazyk /XML - Extensible Mark-up Language/ je nástupcem HTML. Oproti HTML nabízí řadu výhod a základním rozdílem je to, že umožňuje oddělení obsahu a prezentace. W3C zpracovalo desetibodové resumé, shrnující charakteristické rysy XML. Detailnější technické informace najdete v sekci xml webové stránky konsorcia W3C <http://www.w3.org/XML/> nebo na <http://www.xml.com/>. S XML je stále spojena celá řada výzkumných a vývojových aktivit – informace o aktualizaci nejnovějších dotazovacích nástrojů jako např. **xpath** jsou k dispozici na stránce W3C XML Query Group <http://www.w3.org/XML/Query>. Knihovny by měly zvážit také přechod na XHTML. Sdílení metadat a XML je stručně popsáno (včetně omezení a způsobů podpory efektivního sdílení informací v různých souvislostech) ve zprávě organizace UKOLN <http://www.ukoln.ac.uk/nof/support/help/papers/metaxml.htm> z Velké Británie.

RDF

Systém popisu zdrojů /RDF - Resource Description Framework/ umožňuje kódování, výměnu a vícenásobné využití strukturovaných metadat s pomocí XML jako syntaktického nástroje vzájemné výměry dat. Tímto způsobem podporuje také integraci různých aplikací od knihovních katalogů a celosvětových adresářů; až k publikování a třídění zpráv; softwaru a dokumentům; soukromým sbírkám hudebnin, fotografií a dokumentaci různých událostí a akcí.

RDF umožňuje také zpracovávání přehledů a zpráv o zdroji jako souboru vlastností, které odpovídají určitému schématu. Přehledy jsou zaznamenávány ve formátu rdf: Popis elementů XML.

Důvodem, proč je RDF tak rozšířený, je fakt, že zpracovává strukturovaná data, kóduje je a vyměňuje standardizovaná metadata, což umožňuje propojitelnost mezi různými soubory metadat definovaných různými zdroji popisu. RDF kromě toho poskytuje prostředek pro publikování jak okem čitelných tak strojově zpracovatelných slovníků vytvořených k podpoře vícenásobného využití a rozšíření metadat ve společenstvích na různém stupni informační gramotnosti - viz úvod do RDF <http://www.dlib.org/dlib/may98/miller/05miller.html>. Popis RDF často využívá knihoven jako analogie. Technické informace viz webové stránky konsorcia W3C RDF. SeamlessUK je jedním z příkladů místního informačního systému, který využívá ke kódování metadat RDF.

Z39.50

Z39.50 je mezinárodní vyhledávací protokol /ISO 23950, 1998/ umožňující prohledávání /obvykle vzdálených/ různorodých databází a získávání dat na základě jednoho uživatelského rozhraní. Z39.50 definuje standardní způsob komunikace a předávání informací mezi dvěma počítači. Byl vytvořen k podpoře vyhledávání a získávání informací - plnotextových dokumentů, bibliografických dat, obrazů a multimédií - je založen na architektuře typu klient-server a je plně funkční v prostředí Internetu. Umožňuje uživatelům vyhledávat v několika katalogích či jiných databázích v jednom integrovaném vyhledávání.

Poznámka:

Dokud se dotazovací jazyky XML dále nerozvinou, bude Z39.50 stále preferovaným vyhledávacím protokolem pro systémy nabízející komplexní distribuční vyhledávání.

Harvest

Harvestingový software je způsob "sklizně" či shromažďování metadatových informací ze seznamu předem určených webových zdrojů, například webových stránek spolupracujících organizací. V projektu Seamless /viz výše/, vytváří Harvest indexový soubor, který systém prohledává na uživatelský dotaz. Výsledky se pak integrují s výsledky vyhledávání pomocí Z39.50 a jsou uživateli předány v podobě jednoho seznamu. Kliknutím na jakýkoliv odkaz se buď zobrazí webová stránka /pro dokumenty z Harvest/, nebo odkaz na databázi /pro dokumenty z Z39.50/.

Prohledávání prostřednictvím softwaru Harvest tvoří základy aktivit Inicitaivy otevřených archivů <http://www.openarchives.org/> /OAI - Open Archives Initiative/, což je sdružení, jehož cílem je zjednotřit

vyhledávání informací na Internetu pomocí rozvoje a aplikace příslušných univerzálních norem a standardů. Představuje také potenciál pro oblast muzeí.

Vyhledávání obrazových informací

Do nedávné doby musely být obrazové informace vyhledávány pomocí textového popisu či klasifikačních kódů v některých případech podporovaných textovými soubory či přizpůsobenými speciálně k práci s obrazovými informacemi. AAT /Umělecký a architektonický tezaurus/ obsahuje 120.000 termínů pro popis objektů, textových informací, obrázků, architektury a materiálů týkajících se kulturního dědictví. Obrazové informace mohou být také klasifikovány pomocí systémů jako jsou ICONCLASS <http://www.iconoclass.nl/> pro umělecké práce a muzejní výstavy; TELCLASS pro televizi a video a "Klasifikace společenské historie a průmyslu" (Social History and Industrial Classification) pro muzejní exponáty.

Více moderních systémů dokáže v současné době zpřístupňovat obrazové dokumenty, které nebyly verbálně popsány.

CBIR /Vyhledávání obrazových dokumentů na základě obsahu - Content Based Image Retrieval/ nepoužívá indexování pomocí klíčových slov. Obraz se vyhledává na základě charakteristik samotného objektu samotnému, jako jsou barva, struktura či tvar. Například plážová scenérie by byla nahoře modrá a dole žlutá. V současné době probíhá odborná diskuse <http://www.unn.ac.uk/iidr/report.html> týkající se různých typů CBIR. Existují komerční systémy CBIR: IBM má QBIC <http://almaden.ibm.com/>, kde je obraz popsán na základě barvy a tvaru a software vyhledává odpovídající obrazové dokumenty. Jako demonstrace může posloužit webová stránka muzea Ermitáž http://www.hermitagemuseum.org/html_En/07/hm7_41_1.html.

Na téma popisu a klasifikace obrazů existuje diskuse na Curcus (ve francouzštině). Odborná diskuse probíhá také na stránkách Diffuse, která se věnuje "ztrátovým" versus "neztrátovým" způsobům vyhledávání obrazových dokumentů: soubory, které jsou předávány elektronicky a mají odpovídající obrysovou ostrost, jsou objemnější a pomalejší než ty, které tuto ostrost nemají.

ÚKOLY PRO BUDOUCNOST

Sémantický web je více popsán v časopise Scientific American z května 2001. Podle Tima Berners-Lea, ředitele Konsorcia celosvětové sítě /World Wide Web Consortium W3C/: "Sémantický web je síť dat, která svým způsobem připomíná globální databázi. Je založen na myšlence, že "Cílem www je, aby byl využitelný nejen pro komunikaci člověka s člověkem, ale aby také počítače byly schopny komunikovat a pomáhat". Bude ale trvat ještě dlouho, než se tento sen uskuteční. Budou ho postupně vytvářet lidé s různými zájmy. Skutečná síla sémantického webu bude známá ve chvíli, kdy lidé vytvoří celou řadu programů, které budou shromažďovat na síti dokumenty z různých zdrojů, zpracují informace a vymění si vzájemně výsledky.

V této následné fázi se Sémantický web vymaní z virtuální reality a rozšíří se do fyzického světa. Více mikrovlnné trouby schopné prostřednictvím Internetu konzultovat postup vaření přímo s webovou stránkou dodavatele zmražených potravin může být rozšířena na služby poskytované veřejnými knihovnami, archívy a muzei. Virtuální obrazový rámec může například konzultovat zobrazení ideálního obrazu přímo s místním muzeem či uměleckou galerií pomocí jednoduchého hlasového dotazu, stejně jako se informovat o stažení biografie v podobě elektronické knihy v místní virtuální veřejné knihovně. V tomto prostředí by měla být fantazie pracovníků knihovny podporována až na hranici únosnosti.!"

Důležité koncepce, technologie, protokoly a normy pro rozvoj Sémantického webu jsou již vytvořeny anebo se vytvářejí včetně XML, RDF a speciální identifikace.

Ontologie. Ontologie může být definována jako formální popis objektů a jejich vzájemných vztahů. V kontextu Sémantického webu je jejím cílem umožnit přístrojům komunikovat s jinými přístroji s omezenou či nulovou lidskou účastí. Viz Ontoweb a Wonderweb. K ontologiím existují progresivní metody přístupů, zde jsou některé z nich:

- Seznam terminů - nedefinované vztahy
- Klasifikační schémata
- Tezaurus - odkazy a asociované vztahy
- Tématické mapy - nová norma ISO pro systém popisující struktury vědomostí a jejich spojení s informačními zdroji. Měly by poskytnout účinné metody navigování objemných a vzájemně propojených dat. Místo kopírování znaků knižního rejstříku tématická mapa tyto znaky zobecňuje a rozšiřuje současně mnoha různými směry.
- Ontologický jazyk - DAML+Oil - popis logických vztahů

Význam ontologie ve vědomostních systémech řízení a při vývoji Sémantického webu se stále zvyšuje a lze počítat s jejím využitím v takových vědomostních systémech, jako například v elektronickém obchodu pro popis produktů a ve službách či popisu organizací digitalizovaných muzejních sbírek.

Elektronické podpisy, veřejná klíčová infrastruktura pro provádění bezpečných elektronických transakcí a navození důvěry, jsou také považovány za klíčový prvek Sémantického webu. Viz služby na míru.

Webové služby jsou relativně novým pojmem a očekává se, že se budou v průběhu příštích let dále rozvíjet. Mohou být prvním větším praktickým projevem uvažování na základě Sémantického webu. Podrobné definice se různí, ale webové služby umožní rozvoj aplikací softwaru, aniž by bylo nutné vědět, kdo je uživatelem, kde je, ani jiné důvěrné informace o jeho osobě. Viz *Diffuse*. V příštích letech mohou být vyvinuty takové webové služby, kterým se dá lehce porozumět a jež mohou být automaticky využívány uživateli či pracovníky knihovny pomocí jednoduchých přístrojů. Dodavatelé externích aplikačních služeb /ASP/ mohou tyto služby také poskytovat. Webové služby jsou založeny na otevřených internetových normách a standardech.

Základní normy a protokoly pro webové služby se stále vyvíjejí a s jejich dokončením se počítá do konce roku 2002. Obsahují /kromě XML/:

- **WSDL** /Jazyk popisující webové služby - Web Services Description Language/, který umožňuje běžný popis webových služeb
- **UDDI** /Univerzální popis, nalézání a integrace - Universal Description, Discovery & Integration/ registry, které ukazují informace o firmách či jiné skupině a jejich technické profil
- **SOAP** /Jednoduchý objektový přístupový protokol - Simple Object Access Protocol/, který umožňuje výměnu strukturovaných zpráv mezi počítačovými programy.

Koncepce webových služeb je v současné době rozvíjena v rámci reklam (bannerů) elektronického obchodu. Objevují se ale již potenciální aplikace vhodné pro poskytovatele veřejných služeb. Například vyhledávací rozhraní může být zpřístupňováno či poskytováno jako webová služba samotnými veřejnými knihovnami či poskytovateli aplikovaných služeb na zakázku knihovny .

V budoucnu jistě sklídíme ovoce, nespočetných vylepšení na poli vyhledávání informací, díky kterým se současně užívané systémy mohou zdát primitivními. Sofistikovaní osobní "roboti" budou dále systematicky prohledávat Internet, speciálně upravený tak, aby ho mohli „roboti“ snadno prohledávat na základě metadat, řízených a speciálních slovníků. Zobrazí přesně to, co uživatel požadoval, protože budou znát jeho vyhledávací návyky a budou to dělat také v době, kdy budou jinak zaměstnání. Ne zobrazí žádnou zbytečnou informaci a nevynechají nic relevantního bez ohledu na to, kde se dokumenty nachází.

ODKAZY NA WEBOVÉ STRÁNKY

Mezinárodní

Norma IFLA o bibliografickém popisu

http://www.ifla.org/VII/s13/pubs/isbd_m0602.pdf

Národní

Dánsko

BibHit

Katalog internetových zdrojů indexovaných pomocí metadat. Jde o společný projekt tří veřejných knihoven – (Veřejná knihovna v Horo, Ústřední Veřejná knihovna v Aarhusu a Veřejná knihovna v Silkeborgu) - a jedné vědecké knihovny – (Knihovna Aarhuské obchodní školy.
<http://www.bibhit.dk/info/english.htm>

Finsko

Link Library

Adresář internetových zdrojů vybraných a popsanych knihovniky jako společný projekt. Obsahuje přes 6000 odkazů, organizován je finskou verzí Deweyho /ykl/ a je publikován ve dvou jazycích: finsky a švédsky. Obsahuje také vybavení pro metavyhledávání, které vyhledává v knihovních katalozích všech propojených knihoven současně.

<http://www.kirjastot.fi/linkikirjasto/selaus.asp?kieli=suomi&hid=&languageid=>

Německo

Národní katalogizační a indexové schéma

Schéma, které přijaly německé veřejné knihovny pro popis zdrojů: má 25 kategorií. Řada z nich používá centrálně zpracovávaná data, stále častěji poskytovaná online Národní knihovnou či největšími dodavateli jako např.firmou EKZ z Reutligenu.

<http://www.ekz-bibliotheksservice.de/ekz/home.nsf/pagesstartseite>

Nizozemí

OCLC PICA

Paralelní vyhledávání v ústředním nizozemském katalogu a katalozích různých knihoven.

<http://oclc-pica.org/?id=2&ln=uk>

Velká Británie

Co-East, spolupráce několika veřejných knihoven na východě Anglie

<http://www.co-east.net/>

Muzeum v Leedsu - webová stránka nabízející virtuální prohlídku městského muzea

<http://www.leedslearning.net/makingconnections/library/flashcheck.asp?platform=win>

Digitální projekt knihovny v Glasgow

<http://www.gdl.cdli.strath.ac.uk>

Schéma přístupu do knihoven v Sunderlandu

<http://www.lash.sunderland.ac.uk/>

SeamlessUK - distribuovaný systém místních informací

<http://www.seamless-uk.info/>

USA

Solinet

Síť knihoven, jež se snaží zdokonalit přístup uživatelů k informacím a umožnit účastnickým knihovnám efektivně uspokojovat potřeby vzdělávání, ekonomického rozvoje a zlepšení životní úrovně v regionu.

<http://www.solinet.net/index.cfm>