

Kdo hledá, najde – přehled současných možností pro vyhledávání odborných informací

Daniela Tkačková *

Daniela.Tkacikova@vsb.cz

Abstrakt: Příspěvek upozorňuje na některé změny v tradičních způsobech komunikace odborných informací, k nimž dochází pod vlivem technologického rozvoje v posledních deseti letech. Tyto změny přináší otevřenost přístupu k informacím, jež je jedním ze základů všeobecné dostupnosti informací. Díky těmto pozitivním změnám je možné řadu volně přístupných sekundárních informačních zdrojů účinně využívat pro zjišťování a získávání odborných informací; mnohé z těchto zdrojů mohou být *za určitých podmínek* plnohodnotnou alternativou k tradičním finančně náročným bibliografickým bázím dat. Lze je využít jak pro vyhledávání odborných informací uživateli, tak v rámci některých služeb knihoven (např. v MVS pro zjišťování přesných údajů).

Klíčová slova: elektronické publikování, vyhledávání odborných informací, volně přístupné sekundární informační zdroje, alternativní přístup k informacím

1 Úvod

Rozvoj moderních informačních technologií v uplynulém desetiletí nebyvalým způsobem otevřel přístup k odborným informacím. Přestože nejde vždy o skutečně otevřený přístup až k plným textům primárních dokumentů, je možné již dnes, díky otevřenosti přístupu alespoň k sekundárním informacím a díky jejich vzájemnému propojování, bez velké nadsázky říci, že myšlenka všeobecné dostupnosti informací, jež doprovázela i vznik Webu (a která samozřejmě dodnes rozvoj Webu stále provází), začíná nabývat reálné obrysy.

Pojem „všeobecná dostupnost informací“ je spojen s řadou dalších aktivit, a tak pro každého zájemce o informace týkající se předmětu jeho zájmu představuje uskutečnění této myšlenky něco jiného. Mezi uživateli je běžně rozšířenou představou, ovšem stěží uskutečnitelnou, například již s ohledem na autorská práva či oprávněně očekávaný finanční zisk, názor, že všechny informace budou volně dostupné pro kohokoliv, kdo je v daném okamžiku bude potřebovat. Možná k tomu jednou opravdu dojde, ale zatím stačí, že se způsobem, ještě i před deseti lety nejspíš těžko představitelným, rozšířila nabídka volně přístupných zdrojů usnadňujících (nejen) dostupnost odborných informací. I když v trochu jiné rovině, než bylo tvůrci Webu původně zamýšleno, je to právě zásluhou této internetové služby. Nové technologické možnosti doprovázející rozvoj a rozšíření Webu byly předpokladem významných změn ve vydavatelské činnosti, samozřejmě nejen v oblasti orientované na komunikaci odborných informací. Jejich důsledkem je kromě jiného i poměrně snadná dostupnost bibliografických informací, v určitých případech také plných textů publikovaných dokumentů. I kdyby do budoucna zůstalo „jen“ u otevřeného přístupu k bibliografickým informacím, mohlo by se z pohledu uživatelů jednat o naplnění myšlenky všeobecné dostupnosti informací; většina těchto informací totiž byla ještě poměrně nedávno obtížně přístupná, ať už bariérami bránícími v přístupu byly vysoké finanční náklady nebo překážky fyzické.

* Vysoká škola báňská-Technická univerzita Ostrava, Ústřední knihovna, 17. listopadu 15, 708 33 Ostrava-Poruba

S velkým zjednodušením lze říci, že za volným přístupem k informacím stojí na jedné straně snaha vydělat peníze, na druhé pak zbytečně neutráct znovu za to, co již bylo (ne)jednou zapláceno. Předmětem příspěvku není zkoumání příčin tohoto jevu, ani prognóza možného vývoje v budoucnosti; měl by především upozornit na některé volně přístupné informační zdroje, které jsou za určitých podmínek již dnes plnohodnotnou alternativou k jinak finančně nákladným (a tedy mnohdy stále nedostupným) zdrojům a službám. Nicméně neuškodí se přece jenom zastavit u několika zajímavých věcí, které s možnostmi využívání volně přístupných zdrojů souvisejí. Současný stav totiž vyvolává řadu otázek a nutně vede také k přemýšlení nad tím, jakým způsobem budou tyto informace zpřístupňovány v (blízké?) budoucnosti.

2 Elektronické publikování a změny v přístupu k informačním zdrojům

Pokud jde o zdroje, které je dnes možné *volně* využívat pro zjišťování či dokonce získávání odborných informací, jsou to většinou nejen profesionálně vytvářené a zpřístupňované zdroje, ale současně také zdroje komerční. Lze se díky nim dostat ke stejným informacím jako prostřednictvím tradičních finančně velmi náročných bibliografických bází dat. Mnohdy již dnes umožňují přístup dokonce k širšímu okruhu informací, zvláště s ohledem na hloubku retrospektivy (například Royal Society of Chemistry má své archivní časopisy přístupné počínaje rokem 1841, Institute of Physics Publishing od roku 1873 apod.). V minulosti v dobách výhradně tištěných primárních informačních zdrojů se až na výjimky (především vědecké společnosti, viz např. American Chemical Society, Institution of Electrical Engineers apod.) sami nakladatelé nevěnovali zpřístupňování sekundárních informací formou tištěných bibliografií (nebo referátových časopisů) či bibliografických bází dat. Katalogy knižní produkce nakladatelství, stejně jako informace o dalších produktech vydavatelské činnosti, šířené rovněž klasickou cestou, jsou sice určitým typem sekundárních informací, měly však vždy za cíl především prodat vydané publikace. Pokud chtěl běžný uživatel získat odborné informace, nejčastěji se k nim dostal prostřednictvím knihovnických fondů a knihovnických služeb. Mezi nakladateli a knihovnami se ovšem s nárůstem publikovaných informací prosadili také producenti sekundárních informačních zdrojů a distributoři.

Elektronické publikování v posledních deseti letech však mnohé z těchto „ustálených zvyklostí“ narušilo, což nabízí příležitost ke krátkému zamyšlení a k položení několika otázek na téma budoucnosti komerčně šířených bibliografických bází dat. Přestože to zatím nejspíš není na pořadu dne, je namísto celkem přirozeně otázka, jak dlouho asi vydrží dosavadní model zpřístupňování informací, který právě tyto sekundární informační zdroje představují. Jaký typ přidané služby by v dlouhodobé perspektivě musely poskytovat, aby tím vyrovnaly minimální náklady, které uživatele stojí vyhledávání těchto informací ve volně dostupných zdrojích nakladatelů či ve volně přístupných zdrojích financovaných z veřejných prostředků, ve srovnání s náklady na přístup k těmto bázím dat.

Z pohledu běžného uživatele nemůže tento rozdíl vyvážit ani relativní úplnost registrovaných dokumentů, ani například řízený slovník. Pokud jde o rozsah registrovaných informací, jsou zatím tradiční bibliografické báze dat komplexnější než většina volně přístupných zdrojů. Na rozdíl od těchto zdrojů mají báze dat určitou výhodu i v tom, že jde zpravidla o zdroje oborově zaměřené (což může být z hlediska snahy zajistit jejich dostupnost např. v instituci typu univerzity také nevýhodou; čím více oborů, tím více finančně náročných zdrojů se zdá být nezbytných...), zatímco volně dostupné zdroje, samozřejmě v závislosti na rozsahu a míře specializace vydavatelského programu daného nakladatelství, bývají často polytematické. Další nespornou výhodou tradičních bibliografických bází dat je analytické sledování některých zdrojů, především konferenčních materiálů, jež doposud nejsou v takové

šří a hloubce retrospektivy ve volně přístupných zdrojích vyhledatelné. Tím (s určitým zjednodušením) výčet kladů tradičních komerčních bibliografických bází dat můžeme zakončit. Jejich nevýhodou jsou především vysoké finanční nároky, které způsobují, že jsou i dnes pro řadu uživatelů nedostupné. Koncový uživatel ovšem většinou náklady na využívání komerčních bibliografických bází dat přímo nehradí, nicméně ten, kdo rozhoduje o financování přístupu k nim, by současné či budoucí alternativní možnosti zjišťování a získávání odborných informací v úvahu vzít měl.

Rychlost, s jakou významní vydavatelé odborných informací zpřístupnili svoji historickou časopiseckou produkci, opravňuje k domněnce, že po nich bude následovat digitalizace a zpřístupnění dalších uživatelům žádaných publikací. Vzájemné propojování časopiseckých článků prostřednictvím citací a nástroje usnadňující jejich vyhledávání napříč vydavateli, překračuje fyzickou izolovanost jednotlivých nakladatelských sbírek, čímž postupně získávají uživatelé možnost jediným dotazem prohledat podstatně větší rozsah informací, než je tomu u některých tradičních bází dat. Další výhodou je možnost vyhledávání v plných textech dokumentů, bez ohledu na to, zda k nim má uživatel předplacen přístup či nikoliv. To je důležité především proto, že tyto zdroje jsou orientovány na koncové uživatele, kteří si navíc dokážou na základě znalosti hledaného tématu mnohem lépe vyhodnotit potřebnost nabídnutých odkazů než v tradičním modelu vyhledání za pomoci prostředníka, byť tzv. informačního specialisty. Z tohoto modelu se přes všechny inovace stále někteří producenti tradičních bází dat nemohou vymanit. Koneckonců nemají většinou ve svých databázích plné texty registrovaných dokumentů. A pokud je mají, jsou svázáni určitými omezeními ze strany vydavatelů, nebo zpřístupňují i takové zdroje, které svou kvalitou nevyvážejí vysoké náklady na zajištění přístupu k nim.

Zatím se všichni zprostředkovatelé předhánějí v nabídce nových nástrojů pro integraci elektronických zdrojů (ať už plnotextových či bibliografických) mezi tradiční fondy a služby knihoven. Jejich účelem má být nejen jejich snadnější využívání uživateli, ale také jednodušší správa na straně knihovny. Snaha rychle se orientovat v nepřehledném množství nabídek nových zdrojů, nástrojů i služeb s cílem zvolit to nejpříjemnější, nejvhodnější, nejperspektivnější, zbytečně finančně nevyčerpávající..., zkrátka to nejúčinnější řešení, samozřejmě – jak jinak – především s ohledem na uživatele, může vést i k tomu, že se ve skutečnosti ocitneme v některé ze známých cimrmanovských slepých uliček. Troufám si dokonce tvrdit, že leckdo se již v několika z nich ocitl.

3 Výsledky změn ve vydavatelské činnosti

Po období, které je možné s jistým zjednodušením charakterizovat tak, že vydavatelé vydávali a prodávali primární dokumenty knihovnám, jež je zpracovávaly, uchovávaly a zpřístupňovaly uživatelům, od jisté doby díky jinak obtížně zvládnutelnému množství informací rovněž s pomocí producentů sekundárních informačních zdrojů a různých typů distributorských firem, se dnes zásluhou rozvoje moderních informačních technologií ocitáme v situaci, kdy již mezi vydavatelem a uživatelem v řadě případů není de facto potřebný žádný další prostředník. Pochopitelně s výjimkou nezbytných technologických nástrojů, bez nichž elektronické publikování, zpřístupňování i využívání informací není možné (hardware, software, (tele)komunikace...). Pro pracovníky knihoven, ovšem nejen pro ně, je tedy důležité přijmout fakt, že se díky tomu stále větší množství informací dostává od autora k uživateli jinak než tradičními cestami, na něž doposud byli (všichni) zvyklí. Změny se netýkají jen časopiseckých informací, ale také knih, zůstaneme-li v oblasti tradiční produkce. Že je dnes tzv. šedá literatura často snadněji dostupná než knihy či časopisy, alespoň tedy pokud jde o dostupnost plných textů, snad ani není nutné zdůrazňovat.

Pokud jde o vydavatele samotné, týkají se tyto kroky směrem k otevřenosti přístupu k informacím nejen těch, kteří podporují a realizují otevřený přístup (open access) [1] k výsledkům vědecké činnosti i ke kulturnímu dědictví (viz například myšlenky vyjádřené v dokumentech iniciativ jako jsou *Declaration of the Budapest Open Access Initiative*, *Bethesda Statement on Open Access Publishing* nebo *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* [2, 3, 4]), ale samozřejmě a dokonce mnohem více, vezmeme-li v úvahu rozsah zpřístupňovaných zdrojů, také komerčních nakladatelství, jejichž produkce je stále základním zdrojem odborných informací a kteří respektováním osvědčených vydavatelských postupů a pravidel zaručují určitou úroveň kvality publikovaných informací, jejich vyhledatelnost i dlouhodobou (trvalou) dostupnost. Vedle toho rozvoj Internetu přinesl i další možnosti přímo pro samotné autory, kteří mohou „vydávat“ své práce „vlastním nákladem“ podstatně jednodušeji, než tomu bylo v tradiční tištěné komunikaci informací; to je však při publikování elektronickou cestou oblast z řady důvodů problematická, obzvlášť tehdy, když autor ke zveřejnění využije i vlastní technologické zázemí; k možné otázce sporné kvality publikovaných informací se tak přidružuje ještě další problém, jenž souvisí se zajištěním jejich dlouhodobé dostupnosti.

Rozvoj nových technologií umožnil nakladatelům „znovu vydat“ všechny doposud publikované a dávno prodané ročníky časopisů. Za touto aktivitou je možné si představit leccos, celkem jistě je však jednou z důležitých motivací finanční zisk. Moderní informační a komunikační technologie umožňují navíc prodávat na rozdíl od tradičních tištěných forem dokumentů jednotlivé části dokumentů (články, kapitoly knih, jednotlivá hesla v encyklopediích a slovnících apod.), už z tohoto důvodu je nezbytné, aby k informacím o nich vydavatelé umožnili volný přístup.

Vzhledem k aktivitám komerčních nakladatelů, nejen v té v současnosti nejviditelnější a relativně stabilní oblasti, tedy v oblasti zpřístupňování časopiseckých informací, se tedy nezdá být únosné dlouhodobé mnohonásobné zpřístupňování týchž informací. Současně s ohledem na obchodní modely, které umožnily moderní informační technologie, je obtížné si představit, že by nakladatelé po nějaké (zkušební) době opět uzavřeli své informační zdroje a neumožnili uživatelům, svým potenciálním zákazníkům, jejich bezplatné prohledávání a využívání, alespoň na úrovni zpřístupnění sekundárních informací.

V současnosti volně přístupné sekundární a často i primární zdroje jsou tedy vhodnou alternativou k tradičním placeným informačním zdrojům, zcela jistě tehdy, pokud na ně neklademe další nároky, jimž (zatím?) nemohou z řady důvodů vyhovět. Alternativou, na kterou se zatím nelze výhradně spolehnout a která je spíše doplňkem tam, kde nebylo možné pokrýt potenciální potřeby uživatelů zajištěním tradičních zdrojů z nedostatku finančních prostředků. Případně tam, kde se jedná o potřeby spíše okrajové či náhodné. Volně přístupné informační zdroje sice nevyžadují finanční prostředky, je však nezbytná jejich dobrá znalost; je důležité orientovat se v dostupných zdrojích a službách a dokázat posoudit, které z nich se vyplatí nabídnout uživatelům jako doplněk či za určitých okolností jako náhradu tradičních bází dat. Jde většinou o zdroje a služby nakladatelů samotných (zpravidla se jedná o zpřístupnění jejich časopisecké produkce) nebo o služby, které umožňují souhrnný přístup k elektronickým časopisům (jako například Ingenta [5] nebo Extenza [6]), volně přístupné specializované báze dat financované z veřejných prostředků (Medline [7], Eric [8], Agricola [9]), databáze a služby zpřístupňující speciální dokumenty (nebo alespoň informace o nich), například patenty a normy, a šedou literaturu (vysokoškolské kvalifikační práce, výzkumné zprávy, preprinty apod.) [10].

V počátcích rozvoje a využívání Webu se často o knihovnicích hovořilo (lépe řečeno: hovořili tak o sobě sami knihovníci) jako o navigátorech, kteří by měli uživatelům ukazovat

cestu ke kvalitním zdrojům na Internetu. Pokud jde o běžné webové zdroje, těch je dnes tolik, že je vhodnější ponechat jejich vyhledávání a hodnocení kvalitním vyhledávacím službám. Pokud jde o ty ostatní zdroje, které jsou zpravidla více či méně spjaty s tradičním publikováním (stejně jako tradiční služby knihoven), pak právě tady mohou knihovníci uplatnit ony přidané hodnoty, na něž se často odvolávají, tedy své znalosti a zkušenosti získané prací s tradičními zdroji.

Volně přístupné zdroje je možné nabídnout uživatelům pro vyhledávání informací v případě, že knihovna (instituce) nemá přístup ke komerčním bibliografickým či plnotextovým databázím, jsou to však zdroje, které mohou v řadě případů využít i sami pracovníci knihoven. Často je nutné doplnit neúplné citace či opravit nesprávné údaje poskytnuté uživatelem, například pro potřeby meziknihovnických výpůjčních služeb nebo v akvizici. Vyplatí se tedy o takových zdrojích vědět, ať už pro potřeby vlastní práce či v rámci služeb uživatelům, znát jejich výhody i určitá omezení. Mnohé z těchto služeb si dříve či později najdou uživatelé sami, určitě k tomu nepotřebují pomoc knihovníků; jedná se přece o volně přístupné webové zdroje. Běžným uživatelům však schází další znalosti a zkušenosti, jimiž by naopak měli být vybaveni pracovníci knihoven. Patří k nim například schopnost vyhodnotit kvalitu takového zdroje ve vztahu k dalším zdrojům, volně přístupným i placeným, především z hlediska obsahu a šíře (úplnosti) registrovaných informačních zdrojů.

4 Výhody a nevýhody volně přístupných zdrojů

Jednou z výhrad uplatňovaných vůči volně přístupným zdrojům je jejich neúplnost ve vztahu k množství publikovaných informací. Běžný uživatel ovšem zpravidla nepotřebuje mít vyčerpávající přehled o všem, co k danému tématu bylo publikováno. Stačí mu ty zdroje, které je schopen zpracovat. Kromě toho, ačkoliv jsou zdroje jednotlivých nakladatelů omezené, je dobré si připomenout, že v posledních letech došlo (a zřejmě bude docházet i v budoucnu) ke sloučení mnoha původně samostatných nakladatelství (jen namátkou: Elsevier a Academic Press nebo Springer-Verlag a Kluwer Academic Publishers; nejen tato nakladatelství se již předtím zpravidla sloučila s některými „menšími“ vydavateli), že se nakladatelé spojují při budování a zpřístupňování nových elektronických služeb (Blackwell a Munksgaard: Synergy [11]) a že také dochází ke spojování nakladatelů s producenty sekundárních informačních zdrojů (opět může jako příklad posloužit Elsevier). Mnozí menší vydavatelé (samozřejmě nejen oni) zpřístupňují svoji časopiseckou produkci prostřednictvím specializovaných služeb (Ingenta, Extenza). Výsledkem je, že ani roztržitost volně přístupných zdrojů a služeb nepředstavuje nepřekonatelnou překážku či časovou ztrátu při zjišťování informací o tom, co bylo publikováno na dané téma. Současně je však nutné uvést, že k podobnému jevu, tj. ke slučování firem, dochází také v oblasti producentů sekundárních informačních zdrojů a navazujících služeb.

Volně přístupným zdrojům bývá rovněž vyčítáno, že neposkytují dostatečný komfort pro formulaci dotazu (ve srovnání s tradičními bibliografickými bázemi dat). To je ovšem pohled knihovníka, informačního profesionála, který zná možnosti tradičních systémů a ví, jak co nejpřesněji s využitím tradičních technik zformulovat vyhledávací dotaz. Běžný uživatel není takovou představou nijak zatížen, podobnými znalostmi a dovednostmi nedisponuje a nemá zpravidla čas ani zájem je získat; je totiž „zkažen“ relativně jednoduchými vyhledávacími nástroji, které mu zpravidla poskytují i prostřednictvím jednoduchého rozhraní dostatek možností pro vyjádření jeho potřeb. Kromě toho už dávno není pravdou, že volně dostupné zdroje neumožňují využít při vyhledávání podobná kritéria a postupy jako tradiční zdroje.

5 Příklady současných možností zjišťování odborných informací

V posledních letech se projevila především ohromný nárůst počtu vědeckých časopisů dostupných v elektronické formě na Internetu. Rozšířil se okruh vydavatelů zpřístupňujících svoji časopiseckou produkci vedle tištěných verzí také elektronicky, ať už na vlastních serverech nebo prostřednictvím služeb distributorských firem či agregátorů. Vzrostl rovněž počet časopiseckých titulů dostupných elektronicky nejen od poloviny 90. let minulého století, tedy od doby, kdy se tato forma publikování začala v širším měřítku uplatňovat. Mnozí vydavatelé již mají v úplnosti zpřístupněny také archivy svých publikací, nejen s časopiseckými zdroji, ale například i s materiály z konferencí (především vědecké společnosti). Hloubka retrospektivy a tím i rozsah volně přístupných zdrojů pro vyhledávání tak přesahuje mnohonásobně rozsah zdrojů registrovaných v tradičních bibliografických databázích. Současná technologická úroveň navíc usnadňuje vzájemné propojení napříč vydavateli způsobem, který byl v tradiční tištěné komunikaci nemyslitelný. Vyplatí se proto zastavit u jednoho z nástrojů, který toto vzájemné propojení umožňuje.

5.1 Digital Object Identifier (DOI)

Digital Object Identifier [12] je URN (Uniform Resource Name). Je to jedinečný a stálý identifikátor přidělený digitálnímu objektu; identifikuje tedy pouze jedinou entitu. Jedná se o strojem čitelný identifikátor umožňující komunikaci mezi různými systémy. DOI může obsahovat další identifikátor (např. ISBN, ISSN), což je jeden ze znaků jeho interoperability. Kromě řady dalších funkcí je nástrojem, který usnadňuje nakládání s digitálními objekty i z pohledu duševního vlastnictví. Identifikátory DOI jsou dnes běžně přidělovány časopiseckým článkům v jejich elektronických verzích, ale mohou být a jsou využity pro jakékoliv objekty umístěné na síti. Přidělování těchto identifikátorů je svěřeno registračním organizacím, které zastřešuje a řídí International DOI Foundation (IDF) [13]. Využití DOI je mnohostranné. V té nejjednodušší podobě slouží DOI k přesměrování na URL. Jednou z hlavních výhod odkazování na digitální objekt prostřednictvím DOI je stálost tohoto identifikátoru. To umožňuje snadno a bez jakéhokoliv přerušení integrovat objekty, jimž byl DOI přidělen, do nových služeb. Stačí pouze aktualizovat adresy (URL), kdykoliv dojde k přemístění na jiný server. Jako nástroj pro zjišťování aktuálního umístění digitálních objektů je využíván Handle System [14]. Společně s DOI a URL jsou v rejstříku uložena i metadata popisující daný digitální objekt, případně různé formáty, ve kterých je zveřejněn (HTML, PDF, XML) či jeho další umístění na síti. To umožňuje přesměrování uživatele na jím upřednostněný datový formát či server. Využívání DOI je základem vzájemného propojení digitálních objektů prostřednictvím citací v rámci spolupráce nakladatelů, kteří jsou členy systému CrossRef.

5.2 CrossRef System

CrossRef [15] je produktem organizace nakladatelů (Publishers International Linking Association), jejichž spolupráce umožňuje propojení bibliografických odkazů ve vědeckých časopisech s využitím DOI. V polovině dubna 2005 se na spolupráci účastnilo v rámci CrossRef již 1 433 vydavatelů; systém pokrýval 11 628 titulů časopisů a registroval 15 693 226 DOI.

Systém CrossRef je využíván pro propojení citací napříč vydavateli. Vydavatelé, kteří jsou členy tohoto systému, mohou vytvářet přímá vzájemná propojení mezi svými online dokumenty prostřednictvím bibliografických citací, tj. mohou v článcích v jimi vydávaných časopisech odkazovat na citované články publikované i jinými vydavateli, tedy umístěné na jiných serverech. Tyto odkazy mohou být ovšem vytvářeny pouze za podmínky, že se údaje

(odkaz) citovaného článku shodují s údaji uloženými v systému CrossRef. Znamená to tedy, že tato propojení se mohou týkat pouze těch odkazů (článků), jejichž vydavatelé již předtím vložili údaje o své produkci dostupné na síti do systému CrossRef.

CrossRef poskytuje technologickou i obchodní infrastrukturu pro vzájemné propojování; využití DOI usnadňuje správu informačních zdrojů: při jakékoliv změně v umístění digitálního objektu není nutné provádět úpravy (změny adres) v citacích nebo v záznamech v databázích; prostřednictvím jediné smlouvy uzavřené s CrossRef je možné propojení s dalšími členy CrossRef.

Kromě propojení prostřednictvím citací, umožňuje systém CrossRef i přímé vyhledávání. Pro vyhledávání je k dispozici několik nástrojů, které využívají metadata digitálních objektů (časopiseckých článků). Jedním z nich je např. CrossRef Search [16], který pro vyhledávání využívá technologie společnosti Google. Systém je tedy možné využít buď pro hledání informací podle tématu, např. názvů článků nebo slov z názvů článků, z abstraktů a dalších věcných údajů, které jsou součástí metadatového záznamu, nebo podle jmen autorů, názvu zdrojového dokumentu apod. Vyhledávat je samozřejmě možné i podle DOI. Jde zatím o pilotní projekt, kterého se účastní jen menší část (41) členů systému CrossRef (např. American Institute of Physics, Blackwell Publishing, Cambridge University Press, IEEE, Institute of Physics Publishing, Nature Publishing Group, Oldenbourg, Oxford University Press, Springer, Taylor & Francis nebo Wiley). Není bez zajímavosti, že je mezi nimi také jedna česká instituce: Ústav organické chemie a biochemie AV ČR, který je vydavatelem časopisu Collection of Czechoslovak Chemical Communications [17].

Další možností je vyhledávání DOI [18], tj. vyhledání umístění dokumentu na síti, které je (z knihovnického pohledu) výhodné využít například tehdy, pokud nemá uživatel k dispozici úplné údaje o hledaném dokumentu. Systém samozřejmě může vyhledat pouze ty dokumenty, kterým byl identifikátor DOI přidělen, nicméně okruh článků opatřených DOI je dnes již rozsáhlý, neboť většina významných vydavatelů má takřka kompletně digitalizovanou svoji časopiseckou produkci, tj. až k prvním číslům prvního ročníku většiny časopisů. Pokud uživatel zná například název zdrojového dokumentu (časopisu), rok vydání a číslo první strany, může si snadno v systému dohledat další informace. Takové stručné údaje často bývají společně s příjmením prvního autora, příp. číslem ročníku, jedinými údaji uváděnými v bibliografických odkazech. Obdobně neúplné či chybné jsou některé informace (citace) vyhledané prostřednictvím Science Citation Index, který je na využívání údajů v bibliografických citacích v člancích založen. Pro potřeby meziknihovnických výpůjčních služeb jsou stručné údaje z bibliografických citací, někdy navíc třeba s chybami v čísle ročníku apod., často obtížně použitelné. Vyhledávání úplných citací například podle čísla strany je možné v tradičních systémech použít jen výjimečně, a tak snaha dohledat chybějící údaje v běžně dostupných zdrojích často nevede k žádoucímu výsledku, je navíc časově velmi náročná. Pokud je tedy časopis, v němž byl hledaný článek zveřejněn, mezi těmi tituly, které jsou již dostupné pro zjišťování DOI, je vyhledání potřebných informací tímto způsobem tou nejrychlejší cestou.

5.3 Google Scholar

Beta verze nové služby Google Scholar [19] byla uživatelům Internetu zpřístupněna zhruba v polovině listopadu roku 2004. Jde tedy zatím o testovací verzi a očekává se, že koncem roku 2005 budou její funkce rozšířeny o další nabídky. Jde o velmi zajímavou a užitečnou službu umožňující prohledávání různých zdrojů zpřístupňujících odborné informace, včetně recenzovaných příspěvků, vysokoškolských kvalifikačních prací, preprintů, technických zpráv, abstraktů i odborných knižních publikací (zde jsou využívány také údaje z projektu

Open WorldCat OCLC [20] pro zjištění informací o tom, která knihovna vyhledanou knihu vlastní, viz odkaz Library Search u záznamů s označením [BOOK]).

Google Scholar je možné využít pro hledání článků publikovaných řadou akademických nakladatelství nebo vědeckých společností, pro prohledávání repozitářů univerzit, stejně jako pro hledání vědeckých prací umístěných na Webu. Google Scholar využívá stejných nástrojů pro formulaci dotazu jako webový vyhledávač Google, což usnadňuje hledání uživatelům. Kromě toho nabízí i některé další funkce, například pro vyhledávání prací určitého autora (author:pistora). Pracuje také s citacemi dokumentů, na které je v odborné literatuře odkazováno; je-li citovaný dokument dostupný na Webu, je možné jej prostřednictvím Google Scholar získat (vyhledat). Díky strategickému partnerství s významnými nakladateli (ACM, IEEE, Emerald, IoP, Springer, Wiley, Elsevier, PubMed a řada dalších), odkazuje na plné texty článků také na jejich serverech. Plné texty jsou samozřejmě přístupné pouze předplatitelům, ovšem na základě dohody společnosti Google s vydavateli musí být všem uživatelům zpřístupněny alespoň abstrakty.

Pokud autoři plné texty svých prací publikovaných ve vědeckých časopisech zpřístupňují ve volně přístupných repozitářích či v archivech preprintů, je možné vyhledat a získat bezplatně jejich plné texty. Google Scholar nabízí ve výsledcích odkazy vztahující se k jednomu dokumentu umístěnému na několika serverech sdružené do jednoho záznamu. Uživatel si může nechat zobrazit jednotlivé záznamy a rozhodnout se, který odkaz mu nejlépe vyhovuje (může přejít na volně přístupný zdroj nebo na server vydavatele, ví-li že má přístup k plnému textu na základě předplatného apod.).

Při hodnocení vyhledaných výsledků využívá Google Scholar osvědčené funkce webového vyhledávače Google a výsledky seřazuje podle relevance vůči dotazu na základě podobných kritérií jako webový vyhledávač; s ohledem na charakter vědeckých dokumentů přihlíží rovněž k citovanosti článků (nabízí také odkazy na zdroje, jež vyhledaný dokument citují), k významu zdrojové publikace apod. Google Scholar automaticky analyzuje citace v dokumentech a zobrazuje je ve výsledcích jako samostatné odkazy s označením [citation]. Tímto způsobem je možné získat informace i o takových příspěvcích, jež doposud nejsou na Webu dostupné (ani žádné další informace o nich). Pokročilé rozhraní služby nabízí možnost upřesnit formulaci dotazu a omezit dotaz jen na prohledávání určité množiny, např. podle časového hlediska.

5.4 Scirus

Podobným zdrojem orientovaným na odborné informace jako Google Scholar je již delší dobu provozovaný vyhledávací server společnosti Elsevier Scirus [20]. Tento nástroj umožňuje současné prohledávání velkého množství časopiseckých zdrojů. Vedle ScienceDirect prohledává také časopisecké databáze partnerských vydavatelů, jako jsou například American Institute of Physics (Scitation [21]), SIAM, PubMed Central [22], patentové informace, ale kromě toho i archivy preprintů, repozitáře univerzit i vybrané volně přístupné webové stránky s odbornými informacemi.

5.5 Open WorldCat

Společnost Google kromě zmínovaného nástroje Google Scholar vyvíjí i další aktivity, které se týkají tradičních činností knihoven. V prvé řadě je to její partnerství s OCLC, které vedlo ke zpřístupnění určité části zkrácených záznamů (asi dva miliony) z katalogu WorldCat v rámci projektu Open WorldCat Program. Vedle Google se na tomto projektu podílí také společnost Yahoo! Prostřednictvím webových vyhledávacích služeb Google a Yahoo Search [23] je tak možné vyhledávat především knižní publikace; Google těchto informací využívá

také v rámci služby Google Scholar. Kombinací údajů o hledané knize (například údaje ISBN) s frází „find in a library“ je možné rychle získat odkaz na záznam vedoucí do konkrétního knihovního katalogu podle potřeby uživatele (například podle toho, která z knihoven vlastnících hledanou knihu, sídlí nejbližší jeho bydlišti). Je dobré znát limity této služby (jde zatím o pilotní projekt [24]), neboť vedle omezené množiny záznamů převzatých z katalogu (WorldCat obsahuje celkem 58 milionů záznamů), hraje velkou roli i současný rozsah zapojení knihoven do projektu (což ovšem není pro uživatele z České republiky tak důležité). Je možné, že i množiny záznamů, které vyhledávací služby zpřístupňují, nejsou shodné, alespoň tomu nasvědčují rozdílné výsledky hledání. Jedním z důvodů by však mohly být i odlišné postupy, jimiž služby záznamy zpracovávaly pro své vlastní databáze.

V poslední době hodně diskutovanou aktivitou společnosti Google je projekt Google Print [25] zaměřený na online zpřístupnění knižní literatury; jakým směrem se vývoj v této oblasti bude ubírat, ukáží až jeho výsledky zpřístupněné uživatelům. Je to však oblast zasahující ještě více do tradiční sféry knihovních služeb, kromě jiného také tím, že Google nabízí knihovnám možnost účasti na svém projektu.

5.6 Příklady dalších užitečných zdrojů

Rozsah volně přístupných zdrojů a služeb, v nichž lze vyhledávat odborné informace obdobným způsobem jako v tradičních bibliografických bázích dat, je v současnosti velmi široký. V první řadě jsou to nakladatelské zdroje zpřístupňující časopiseckou produkci. Vedle prohlížení jednotlivých časopisů postupným procházením ročníků až k jednotlivým článkům, je dnes již standardní službou vyhledávání článků jak prostřednictvím jednoduchého, tak pokročilého rozhraní. Služby jako Ingenta nebo Extenza umožňují rovněž vyhledávání článků z časopisů; jejich výhodou je prohledávání mnohem rozsáhlejší databáze zahrnující časopiseckou produkci více nakladatelství; jediným dotazem je tedy možné dospět k přibližně stejným výsledkům jako postupným prohledáváním služeb jednotlivých nakladatelů.

Vědecké společnosti mají vedle časopiseckých informací zpřístupněny i konferenční materiály, reporty a další odbornou literaturu. Příkladem může být například SPIN Web Database [26], nyní zpřístupňovaná v rámci rozhraní již zmíněné služby Scitation AIP [20], databáze zahrnující asi 1 579 000 záznamů dokumentů z oblasti fyzikálních věd od roku 1975 do současnosti, nebo SPIE Digital Library [27], jež obsahuje více než 175 000 příspěvků z časopisů a konferenčních sborníků SPIE (The International Society for Optical Engineering) od roku 1990 do současnosti.

6 Závěr

V současnosti je obtížné předvídat, jakým směrem se budou možnosti pro zjišťování odborných informací ubírat. To samozřejmě ukáže až budoucnost. Kvalitních nástrojů poskytujících nejen relativně jednoduché nástroje vhodné pro koncové uživatele, ale současně pokrývajících významný okruh kvalitních informačních zdrojů, je dnes velké množství. Je nepochybné, že mohou být s úspěchem využity pro vyhledávání, jak uživateli, tak pracovníky knihoven. V příspěvku bylo možné odkázat jen na některé příklady. Stačí však sednout k počítači, otevřít si třeba vyhledávač Google a začít hledat další zdroje, takové, které nejvíce odpovídají například svou tematikou potřebám uživatelů dané instituce.

WWW odkazy

1. Directory of Open Access Journals, <http://www.doaj.org/>
2. Budapest Open Access Initiative, <http://www.soros.org/openaccess/index.shtml>

3. Summary of the April 11, 2003, Meeting on Open Access Publishing, <http://www.biomedcentral.com/openaccess/bethesda/>
4. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>
5. Ingenta, <http://www.ingentaconnect.com/>
6. Extenza, <http://www.extenza-eps.com/extenza/researcherHome.do>
7. Medline, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>
8. Eric, <http://www.eric.ed.gov/>
9. Agricola, <http://agricola.nal.usda.gov/>
10. OAister, <http://oaister.umdl.umich.edu/o/oaister/>
11. Synergy, <http://www.blackwell-synergy.com/journallist.asp>
12. DOI, The DOI Handbook, http://www.doi.org/handbook_2000/DOIHandbook-v4-2.pdf
13. International DOI Foundation, <http://www.doi.org/>
14. Handle System, <http://www.handle.net/>
15. CrossRef, <http://www.crossref.org/>
16. CrossRef Search (příklad), http://www.iop.org/EJ/search_crossref
17. Collection of Czechoslovak Chemical Communications, <http://cccc.uochb.cas.cz/>
18. DOI Lookup, http://www.crossref.org/02publishers/37guest_query.html
19. Google Scholar, <http://scholar.google.com/>
20. Scirus, <http://www.scirus.com/>
21. Scitation, <http://scitation.aip.org/>
22. PubMed Central, <http://www.pubmedcentral.nih.gov/>
23. Yahoo! search, <http://search.yahoo.com/>
24. Quick facts about the Open WorldCat program, <http://www.oclc.org/worldcat/open/facts/default.htm>
25. Google Print, <http://print.google.com/>
26. SPIN Web Database, <http://scitation.aip.org/spinweb/>
27. SPIE Digital Library, <http://www.spiedl.org/>