

K otázce pokrytí publikační aktivity českých vysokých škol v bibliografických bázích dat

Jaroslav Šilhánek

Vysoká škola chemicko-technologická v Praze

silhanek@vscht.cz

INFORUM 2008: 14. konference o profesionálních informačních zdrojích
Praha. 28. – 30.5.2008

Abstrakt: *Otázky hodnocení výsledků vědecké práce na základě publikační aktivity se dostaly v poslední době díky mimořádnému mediálnímu zájmu do středu pozornosti nejširší veřejnosti a vyvolávají pochopitelně pozitivní i negativní reakce. Základním problémem je především co nejširší a nějakým způsobem charakterizovaná evidence publikační aktivity jako taková, která je pak podkladem hodnocení sama o sobě nebo je na tomto základě vytvářeno hodnocení další. Existence více možných zdrojů pro zjištění počtu publikací pak vyvolává otázky jejich vhodnosti či úplnosti, což je pak prvním zdrojem nedorozumění a pochybností. Na příkladu vysoké školy, jejíž publikační rozsah není sice největší, ale přepočten na počet studentů ji staví na přední místo, jsou demonstrovány rozdíly v evidenci publikační aktivity u dvou hlavních univerzálních bibliografických bází a stručně posouzeny i možnosti jiných zdrojů. Na základě těchto výsledků je pak možné vyslovit závěry o rozsahu a spolehlivosti údajů o publikační aktivitě alespoň v případě určitého oboru.*

Všeobecné poznámky

Otázky hodnocení vědecké činnosti vždy zajímaly jak vědce samotné, tak i ty, kteří jejich činnost financovali nebo rozhodovali o kariéřním postupu i dalších aspektech. Z řady možností se celkem logicky nabízí ta nejvíce viditelná a všeobecně dostupná, a to je publikační činnost, neboli počty a případně i „kvalita“ publikací především ve vědeckých časopisech. To není samozřejmě nic nového a tato praxe probíhá v té či oné podobě dlouhá léta a bez ohledu na více či méně oprávněné připomínky o její ne stoprocentně vypovídající hodnotě je všeobecně akceptována jako v jistém smyslu standard hodnocení každého vědce.

Bylo a je tomu tak i na českých výzkumných institucích. Ovšem otevření se světu a daleko snazší možnosti vzájemné informovanosti i kontroly a především otevřené soupeření o finanční prostředky jako jeden z nejdůležitějších předpokladů vědecké činnosti, způsobily mnohem větší zájem o otázky hodnocení vědecké práce než tam, kde podobné problémy byly již v minulosti tak či onak prodiskutovány. Není proto překvapující, že zveřejnění přehledů hodnocení univerzit a vysokých škol v roce 2007 založené především na publikační aktivitě jejich učitelů [1], [2], způsobilo jisté pozdvižení a jak pozitivní, tak i negativní reakce [3]. Budiž ale na tomto místě velmi důrazně konstatováno, že v žádném případě není důvodem

tohoto referátu se zveřejněnými výsledky a hodnocením polemizovat nebo jej dokonce zpochybňovat. Výsledky jsou nepochybně profesionální a byly získány stejnou a tudíž navzájem srovnatelnou metodikou a v žádném případě je nechceme zpochybňovat nebo jakkoliv doplňovat nebo dokonce žádat opravy. Tento referát je především reakcí na evidentní nedorozumění nebo malou zkušenost s daty o publikační aktivitě, které jsou sice v zásadě jednoduché, ale přece jenom vyžadují, aby jim bylo rozuměno, resp. aby si byl každý vědom jejich významu a možnostem. Výsledkem by pak mělo být alespoň omezení zbytečných sporů a diskusí a větší míra porozumění hodnocení, statistikám nebo pořadí, které jsou a budou publikovány v různé míře i nadále.

Ujasněme si předem některé skutečnosti a různé aspekty zájmu o pokrytí publikační aktivity českých vědců. Možnost dozvědět se co neúplnější odkazy na publikované výsledky vědecko-výzkumné práce celkově nebo za určité období slouží pochopitelně především k vlastnímu šíření vědeckých poznatků a jejich uchování a teprve pak k dalším statistickým nebo evaluačním účelům. Zatímco v prvním případě jde o obecný požadavek zhodnocení prostředků investovaných do vědecko-výzkumné práce, ve druhém případě pak může otázka šíře a případné úplnosti pokrytí takové činnosti způsobovat nesprávné nebo tak či onak zavádějící závěry, což by bylo zvláště závažné, pokud by takové závěry sloužily k dalšímu rozhodování prakticky o čemkoliv. V tomto ohledu ale nemusí jít o „úplnost“ v pravém slova smyslu, ale spíše o pokud možno přesné porozumění, jaké pokrytí daný zdroj nabízí nebo zajišťuje a jaké závěry je možné z nalezených informací dělat.

Konečně je nutné si uvědomit, že jedna věc je možnost dozvídat se o objemu zveřejňovaných výsledků nezávisle, tedy z veřejně přístupných sekundárních informačních zdrojů a druhá věc mít k dispozici údaje poskytované přímo autory. Z logiky celé problematiky je zřejmé, že jakékoliv skutečné hodnocení by mělo vycházet z údajů poskytovaných samotnými autory, případně podle předem jasně stanovených pravidel co „se počítá“ a co nikoliv, se samozřejmou podmínkou ověřitelnosti takových dat. Využívání nezávislých veřejných zdrojů může sloužit k jakýmkoli hodnocení za podmínky, že jsou akceptována celková zaměření a omezení takových zdrojů a že je zásadně dodržována stejná principiální metodika.

Výchozí situace

Konkrétní motivací pro bližší pohled na problematiku hodnocení publikační aktivity vysokých škol, byl evidentní rozdíl ve výsledcích mezi uveřejněnými daty a hodnotami, které jsme si mohli zjistit sami, a to jak ve veřejně dostupných zdrojích, tak i z vlastních interních, ale tak či onak veřejně dostupných zdrojů. Následující tabulka uvádí sumarizaci údajů o publikační aktivitě VŠCHT Praha za rok 2006:

Web of Science		Scopus	
Berger [1]	vlastní údaje	Berger [2]	vlastní údaje
287	342	391	454

Na tomto místě znovu zdůrazňujeme, že zásadně nepochybňujeme publikovaná data a jsme přesvědčeni, že byla získána podle určité metodiky s maximální pečlivostí. Otázkou, která nás zajímá a může být zajímavá i pro širší veřejnost je, proč došlo k relativně větším rozdílům mezi údaji, které uvádí prof. Berger a které jsme zjistili sami hlavně v případě báze dat WoS a v druhém případě, kde jsou rozdíly podstatně menší, proč se liší počty publikací uváděné v bázi dat WoS a Scopus.

Použité a použitelné zdroje dat

Bude užitečné se stručně zmínit o problematice zdrojů, ze kterých je možné informace o publikační aktivitě získávat. V zásadě tak může být využita prakticky každá sekundární bibliografická báze dat, která zahrnuje údaj o roku zveřejnění, protože pochopitelně klíčový údaj, jméno autora, je samozřejmý. Jméno autora je ale v případě snahy získat údaje o větší instituci ne příliš praktické, protože by znamenalo hledat jména i několika set nebo i tisíce autorů. Budiž ale řečeno, že by to bylo nesprávnější a pokud by byl skutečný vážný zájem, existují možnosti takovou cestu volit, i když bude buď pracná nebo, v případě využití služeb databázových center, značně nákladná.

Dalším, v úvahu přicházejícím údajem, je adresa pracoviště autora, resp. autorů. To je cesta, která je nejschůdnější a zřejmě byla použita i ve výše uvedených případech. Má ale dva důležité aspekty. Ten první spočívá v tom, že adresa autora nebo autorů slouží především pro možnost navázat s autorem nebo autory kontakt za účelem diskuse, kladení dotazů nebo požádání o separát práce a za tím účelem stačí adresa jednoho autora, na což je zpravidla zřetelně upozorňováno necht' už role autora s adresou je jakákoliv. Důsledek pak ale je, že adresy ostatních autorů se v bibliografické citaci neobjeví a jejich práce pak logicky nejsou zahrnuty do nalezených počtů. Budiž konstatováno, že toto je naprosto převažující situace, nebo to alespoň převažující situace v bibliografických bázích dat byla. Pro uvádění adres dalších autorů nebyl důvod a v případě tištěných zdrojů by navíc zabíraly místo.

Prakticky existují jen dva velké bibliografické zdroje, kde jsou systematicky uváděny adresy ve smyslu adresy pracoviště, všech na práci uvedených autorů. Je to především Web of Science, kde adresa představuje prvek umožňující vytváření statistických souborů, což je nepochybně dnes jedním důležitých marketingových argumentů tohoto zdroje. A druhým je nedávno na trh uvedená báze Scopus, kde uvádění všech adres je nepochybně součástí možné konkurenceschopnosti právě vůči bázi dat Web of Science. Naprostá většina ostatních bibliografických bází dat adresy všech autorů neuvádí a informace o publikační aktivitě instituce je možné získat jen prostřednictvím jmen autorů, což u menších institucí nemusí být problém, u velkých pochopitelně to problém je.

Problematika identifikace pracoviště autorů v bázi dat Web of Science a Scopus

Jako nejpravděpodobnější důvod výše uvedené diskrepance mezi počty publikací pro VŠCHT Praha v bázi dat Web of Science přichází logicky v úvahu právě problém adresy. Pokusili jsem se proto nalézt pokud možno všechny alternativy adres, které se pro naši vysokou školu

v bázi dat vyskytují, samozřejmě při respektování pravidel, které tato báze má. Nutno konstatovat, že již předchůdce elektronické bázi dat Web of Science, referátový časopis Science Citation Index, zavedl jako jeden z mála referátových zdrojů, poměrně přesná pravidla pro uvádění adres, resp. pro jejich zkracování. Bylo to pochopitelným důsledkem zájmu o maximální snížení pracnosti převádění údajů z tištěných periodik. Používané a jediné dovolené zkratky jsou zobrazitelné, ovšem ne na všechny adresy jsou uvedené standardní zkratky použitelné. Navíc vývoj nabídl některé nové prvky adresy, které se ukazují být jako velice užitečné, konkrétně jsou to všeobecně rozšířená poštovní směrovací čísla a kódy. Výsledek této snahy lze sumarizovat v následující tabulce:

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Inst Chem Technol SAME Prague	236	214	236	221	223	286	266	329	320	332
Prague Inst Chem Technol	66	57	71	65	59	57	51	60	89	81
CZ-16628	25	29	22	27	24	31	35	38	39	33
ICT SAME (Prague OR Praha)	5	5	7	6	7	14	12	8	7	10
VSCHT	6	2	1	0	2	5	2	5	3	2
Vysoka skola chem technol	2	1	3	4	3	3	3	7	7	0
Kombinace všech alternativ	250	222	245	229	234	304	280	340	329	342

Seznam v úvahu přicházejících alternativ adres lze získat buď trpělivých zkoušením nebo využitím možnosti statistického rozboru publikací získaných např. prostřednictvím nejčastěji publikujících autorů. Už z tabulky je vidět, že variací adres je více, ale frekvence jejich výskytu je poměrně malá a zdaleka největší podíl citací má jednu nejčastěji užívanou adresu, nebo její variantu. Příspěvek ostatních alternativ adres představuje cca 3% celkového počtu citací.

Obdobným způsobem byly vyhledány odkazy na publikace autorů z VSCHT v bázi dat Scopus. Je možné konstatovat, že jelikož v tomto případě producent nebyl vázán dlouhodobou tradicí a báze byla koncipována v prostředí velmi pokročilých informačních technologií, mohl pracovat s plnými adresami a zakomponovat i jednodušší postupy zahrnování podobných adres, změn pořadí slov apod. [6].

	2004	2005	2006
Institute of Chemical Technology Prague	208	308	416
Prague Institute of Chemical Technology	208	308	416
166 28	220	257	350
ICT Prague	8	12	16
VSCHT	5	6	10
Vysoká škola chemicko technologická Praha	14	27	28
	299	349	454

Jak je vidět, systém je schopen automaticky registrovat změny pořadí složek názvu. Obdobně jako v případě báze dat Web of Science je výskyt zkratkových adres malý a česká verze názvu je sice rovněž málo frekventovaná, ale nikoliv zanedbatelná. Zajímavý je velmi častý výskyt

směrovacího čísla, výrazně vyšší než v bázi Web of Science. Jinak je rozložení výskytu alternativních názvů velmi podobné v obou zdrojích.

Podstatně složitější než identifikace adresy resp. názvu instituce, je přiřazení hierarchicky níže stojící složky, v případě univerzity fakulta nebo katedra či ústav, v případě Akademie ten který ústav nebo jiné podobné organizační jednotky. Nejenom že je zde ještě větší pestrost buď alternativních nebo ještě detailněji organizačně strukturovaných označení, ale hlavně při převodu do dalších jazyků panuje nejednotnost termínů i na jedné instituci. I když naše kolísání mezi katedrou a ústavem se ve světových bázích asi neprojeví, už německé kolísání mezi *Lehrstuhl*, *Institut*, *Abteilung* ilustruje problematiku. V angličtině naprosto převažující *Department* situaci naopak zjednodušuje. Připustíme-li, že vyhledání údajů na úrovni katedry lze řešit prostřednictvím autorů, tedy členů katedry, zůstává jako skutečně problematická záležitost identifikace fakulty nebo velkého ústavu, např. ústavu Akademie věd.

V našem případě jsme tento problém byli nuceni řešit, protože v publikovaných statistikách jednotlivé fakulty VŠCHT nebyly uvedeny, přičemž se ukázaly se následující skutečnosti. V žádném případě jsme se neseťkali s případem, kdy by v citaci nebyla uvedena instituce, tedy vysoká škola a práce by byla přiřazena jen fakultě nebo katedře či ústavu. V praxi je tedy daleko větší tendence vynechávat název níže stojící organizační složky než naopak. V tištěných verzích bylo zcela logické, že uvádění dlouhých adres včetně fakult nebo i kateder zabírá místo, v elektronických verzích je to už spíše považováno za zbytečnost. Rozborem celého souboru citací pro rok 2006 z Web of Science byly získány následující výsledky:

Počet publikací s uvedenou identifikací ústavu	248
Počet publikací bez identifikace ústavu	94
Celkem	342

Ukazuje se tedy, že v našem případě 27,5 %, tedy číslo blízké se jedné třetině všech citací, nemělo uvedenou nižší složku, fakultu nebo katedru či ústav. Pokud tedy budou na konkrétních institucích požadovány údaje pro vnitřní hodnocení publikační aktivity, může to hlavně u skutečně velkých univerzit nebo institucí představovat nezanedbatelný problém, který ale nemůže být připsán nebo jehož řešení nelze očekávat od žádné sekundární báze dat.

Porovnání výsledků z bází Web of Science a Scopus

Z výše uvedených dat je patrné, že existují dosti výrazné rozdíly mezi počtem publikací nalezených v té či oné bázi dat na základě předpokládaných alternativních adres, přičemž jsme přesvědčeni, že otázka přesného uvádění adres byla jednak pečlivě zohledněna a jednak jsme zjistili, že díky vyhledávacím algoritmům není zas až tak citlivá. Za tohoto předpokladu je pak možné hledat důvody zjištěných rozdílů. Primární počty citovaných publikací za poslední tři roky sumarizované v prvních dvou sloupcích následující tabulky jsou sice rozdílné, ale při prvním posouzení jsou až překvapivě podobné a naznačují na základě propagačních materiálů o bázi dat Scopus očekávaný větší počet citací právě v této bázi [7].

	Scopus	WoS	Scopus-konference	WoS konference	Scopus bez konferencí	WoS bez konferencí	Rozdíly
2004	299	340	53	8	246	332	86
2005	345	329	84	6	265	323	58
2006	454	342	130	13	324	329	5

Obě báze ale ve svých současných podobách nabízejí bezprostřední informaci o typovém složení citovaných dokumentů, kde je jako jedna z forem uváděny odkazy na konferenční materiály. Ty jsou u Scopusu velmi výrazně vyšší než u báze dat Web of Science, což je ovšem obecně známý rozdíl mezi oběma bázemi. Označení typů dokumentů je v případě obou bází jen mírně odlišné, ale liší se prakticky jen v jedné položce, a to jsou právě tzv. kongresové materiály, neboli nejčastěji sborníky vědeckých konferencí.

Articles	296
Conference materials	130
Review	23
Errata	2
Short survey	3
	454

Articles	329
Meeting abstracts	8
Corrections	1
Letter	1
Bibliographic	2
Editorial	1
	342

Jestliže tento typ publikací od nalezených dat odečteme, dostaneme „čisté“ vědecké články, jejichž počet je v případě báze dat Scopus v dřívějších letech výrazně menší než v bázi dat Web of Science, ale už v roce 2006 je prakticky stejný.

Ještě názornější informaci o shodě nebo rozdílech mezi oběma bázemi dostaneme, porovnáme-li seznam titulů vědeckých časopisů a počty článků, které v nich byly nalezeny pro všechny alternativní adresy. Úplný seznam titulů časopisů je pro oba zdroje překvapivě prakticky shodný, ve Web of Science je uvedeno 174 časopisů, ve kterých jsou publikovány články autorů z VŠCHT a za stejné období má stejný seznam z báze Scopus 169 titulů přičemž jmenný seznam je prakticky totožný. Následující tabulka demonstruje na několika prvních časopisech tuto shodu a alespoň v této části prakticky stejné pořadí demonstruje i shodu v počtech článků v jednotlivých případech.

Na několika případech jsou demonstrovány i rozdíly, které ale nespočívají v tom, že by nějaký časopis nebyl zahrnut, ale v individuálních rozdílech v počtu článků. Rozdíly v počtech článků jsou ale opět velmi malé a odchylky se vyskytují na obě strany, tj. více záznamů ve Web of Science nebo více záznamů ve Scopusu, které po sumarizaci za celý soubor dají stejně malý celkový rozdíl pro rok 2006 uvedený v předchozí celkové tabulce. Pro starší období podobný rozdíl nebyl studován, je evidentní, že především v případě báze Scopus je patrný pokračující vývoj a postupné sblížování excerpčních postupů i rozšiřování spektra titulů.

	Source Title	Web of Science	Scopus
1	JOURNAL OF RADIOANALYTICAL AND NUCLEAR CHEMISTRY	25	25
2	CHEMICKE LISTY	14	14
3	CERAMICS-SILIKATY	13	13
4	CZECH JOURNAL OF FOOD SCIENCES	12	12
5	COLLECTION OF CZECHOSLOVAK CHEMICAL COMMUNICATIONS	10	10
6	JOURNAL OF FOOD ENGINEERING	9	9
7	JOURNAL OF CHEMICAL THERMODYNAMICS	7	7
8	CHEMICAL ENGINEERING SCIENCE	6	6
9	DESALINATION	5	5
10	JOURNAL OF THE EUROPEAN CERAMIC SOCIETY	5	5
11	CZECHOSLOVAK JOURNAL OF PHYSICS	4	5
12	FOOD ADDITIVES AND CONTAMINANTS	4	4
13	JOURNAL OF NON-CRYSTALLINE SOLIDS	4	4
14	THIN SOLID FILMS	4	4

15	BIOPOLYMERS	3	3
17	EUROPEAN JOURNAL OF ORGANIC CHEMISTRY	3	3
18	JOURNAL OF CHROMATOGRAPHY A	3	3
19	JOURNAL OF MOLECULAR STRUCTURE	3	2
20	JOURNAL OF SEPARATION SCIENCE	3	3
21	JOURNAL OF SOLID STATE CHEMISTRY	3	3
23	POLYMER	3	3
24	REACTION KINETICS AND CATALYSIS LETTERS	3	3
25	TETRAHEDRON	3	3
26	ANALYTICA CHIMICA ACTA	2	2
28	CHEMICAL ENGINEERING JOURNAL	2	2
29	COMPUTERS & CHEMICAL ENGINEERING	2	2
30	ELECTROCHIMICA ACTA	2	2
31	EUROPEAN JOURNAL OF LIPID SCIENCE AND TECHNOLOGY	2	2
32	EUROPEAN JOURNAL OF NUCLEAR MEDICINE AND MOLECULAR IMAGING	2	0
33	FLUID PHASE EQUILIBRIA	2	2
34	FOLIA MICROBIOLOGICA	2	2

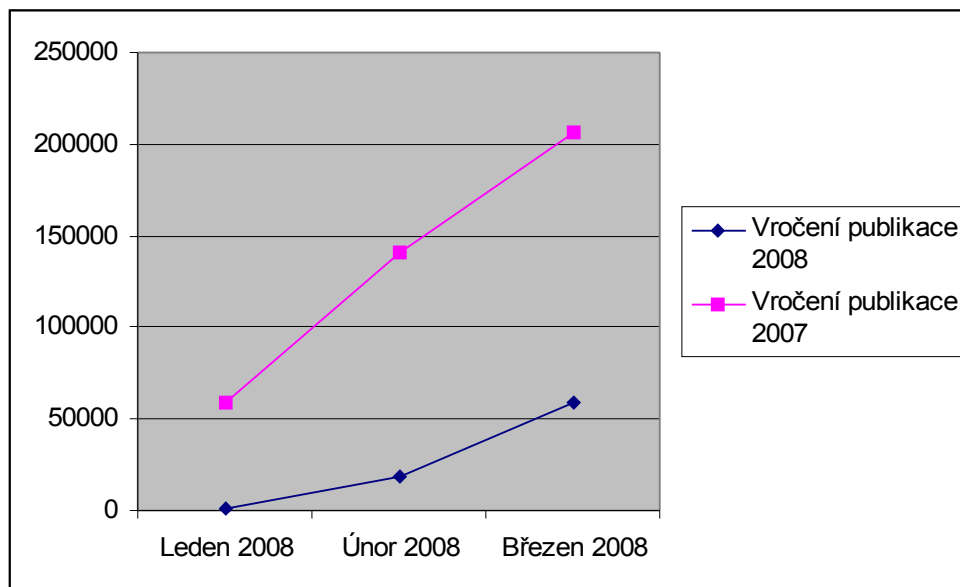
Podrobnější inspekce rozdílů u několika titulů ukázala, že důvod je zřejmě v individuálním posuzování typu konkrétního článku v případě, že se nejedná o vědecký článek v typickém pojetí. Výše uvedené rozdílné počty (+/- jeden článek) jsou způsobeny např. existencí čísla časopisu, který je současně sborníkem konference, přičemž jedna báze zařadí články jako konferenční materiál a druhá jako normální publikace. K podobným rozdílům vedou nejasné materiály v některých časopisech nebo jednotlivých číslech, kdy není zcela zřejmé, zda se jedná o plný text nebo jen část či delší abstrakt přednášky. V každém případě je možné konstatovat, že nebyly shledány prakticky žádné výrazné rozdíly v pokrytí publikační aktivity v případě typických vědeckých článků mezi bázemi Web of Science a Scopus a případné rozdíly jsou celkem snadno vysvětlitelné celkovou koncepcí té či oné báze dat [4], [5].

Další možné zdroje problémů při sledování institucionální publikační aktivity

Rozdílné údaje o publikační aktivitě instituci získané různými subjekty mohou mít ještě jiné důvody než jen pečlivost při zadávání alternativních verzí adres. Je pochopitelné, že zatímco autor může na konci kalendářního roku nebo bezprostředně na počátku roku následujícího přesně uvést, jaké publikace s jeho autorskou účastí vyšly nebo ještě s vročením končícího roku vyjdou, z bibliografickýchází dat je taková možnost teprve tehdy, až jsou zpracovány všechny zdroje patřící do uplynulého roku. To závisí jak na vydavateli, který je schopen distribuovat poslední čísla periodik uplynulého roku ještě během prosince nebo až v lednu roku následujícího nebo i později, tak i na producentu báze dat, jak rychle je schopen data za uplynulý rok zpracovat. Je logické, že to určitou dobu trvá. Představu o tom je možné získat např. sledováním přírůstků v bázi od počátku roku v pravidelných intervalech nebo po známých aktualizacích báze. Jelikož jak u báze Web of Science, tak i báze Scopus nejsou aktualizace jednoduše zřejmé, použili jsme pro kvantifikaci průběhu zpracovávání citací bázi dat Chemical Abstracts, kde ve verzi na CD-ROM je možné vždy zpracovat přírůstek za jeden měsíc. Následující tabulka demonstruje postupné doplňování odkazů s vročením uplynulého a pochopitelně i běžného roku během několika prvních měsíců kalendářního roku.

	Leden 2008	Únor 2008	Březen 2008
Vročení publikace 2008	564	19094	58644
Vročení publikace 2007	58842	141014	206023

Stálý nárůst odkazů na publikace uplynulého roku je velmi markantní, což ještě názorněji demonstruje zobrazení dat formou grafu:



Jak je vidět počet publikací, které by měly být vykazovány jako publikace v uplynulém roce během prvních tří měsíců stále velmi výrazně stoupá a názorně tak demonstruje, že pro nezávislé získávání dat o publikační aktivitě je nutné vyčkat dokončení zpracovávání zdrojů, což přichází v úvahu ne dříve než před polovinou běžného roku. Tedy nejdříve v květnu, ale spolehlivěji v červnu nebo ještě později. Na tuto sice triviální okolnost se někdy zapomíná a výsledky pak zbytečně vyvolávají rozpaky a případnou nespokojenost.

Další možnosti vyhledávání informací o publikační aktivitě instituce

Elektronické formy vědeckých článků otevírají další možnosti zjišťovat jaké publikace byly vydány pod adresou té či oné instituce, např. univerzity, tedy adresy spojené se jménem autora. Především to jsou soubory plných textů článků, resp. elektronické verze vědeckých časopisů, dnes typicky zpřístupňované jako ucelené systémy především velkých vědeckých nakladatelství. Ekvivalence v podobě daného článku mezi jeho tištěnou a digitalizovanou verzí znamená i ekvivalenci adres autora a všech autorů tak, jak bylo vždy dobrou zásadou v oblasti vědeckého publikování. Dnes nejsou žádná omezení co se týče velikosti souborů nebo kapacity paměti a není sebemenší důvod tyto informace omezovat. Stále ale nejsou

adresy všech autorů přebírány z plných textů článků do jejich abstraktů v sekundárních referátových zdrojích přesto, že ty jsou dnes prakticky výhradně v elektronické podobě. Adresy institucí mohou být v souborech plných textů periodik vyhledávány jako textový řetězec a závisí jen na tom, zda příslušný nástroj uvádí adresy do adresovatelných polí, které nabídne zpravidla při volbě „Advanced Search“.

Typickou ukázkou těchto nových možností je např. systém Science Direct nakladatelství Elsevier nebo obdobný nástroj vydavatelství Wiley-Interscience. Tuto možnost zatím nemá systém Springer Link. Pro ilustraci uvádíme výsledky vyhledání publikací VŠCHT Praha v časopisech nakladatelství Elsevier a Wiley-Interscience.

V systému Science Direct je možné vyhledávat pomocí Advanced Search (nebo i Expert Search“) a pole *Affiliation* s volbou příslušného roku vydání. Systém nabízí i možnost kombinace výsledků v Search History, pokud nestačí využití nabízeného formuláře. Pro VŠCHT Praha bylo tak zjištěné, že v roce 2006 bylo v časopisech tohoto nakladatelství publikováno 104 článků.

V souboru plných textů nakladatelství Wiley-Interscience je obdobně možné vyhledávat publikační aktivity instituce v poli *Authors affiliation* spolu s časovým omezením kalendářního roku. Pro náš případ bylo nalezeno 51 článků publikovaných v časopisech tohoto vydavatele v roce 2006. V obou případech se ukázalo jako velice výhodné volit jako vyhledávací složku adresy poštovní směrovací číslo, které poskytovalo vždy největší počet hitů a pomocí formuláře pro Advanced Search a operátoru OR bylo snadno možné kombinovat adresní údaje pro získání relevantních údajů. Celkem tedy jenom u těchto dvou důležitých vydavatelství jsme publikovali 158 článků, což je téměř polovina celkového počtu 342. I toto číslo poskytuje zajímavou informaci z řady aspektů. S největší pravděpodobností by zde měla odpadnout nejistota mezi datem publikace a datem zařazení do sekundární báze, protože se jedná o tituly, které mají z hlediska celého souboru jednoho nakladatelství jasné vřočení. Dnes celkem bezproblémové nabídky exportů nalezených souborů do další nástrojů, např. Excelu, pak umožňuje další operace, slučování, vyhledávání duplicit, kontrolu úplnosti, porovnávání se sekundárními zdroji apod. Lze očekávat, že další vývoj se bude ubírat právě touto cestou, protože plný text vědeckého článku tak, jak je publikován v periodiku, obsahuje nejúplnější afilaci autorů včetně jednoznačného přiřazení každého autora k té které adrese, což se při převodu do sekundárních bází někdy ztrácí.

Závěry:

Otázka pokrytí publikační aktivity českých vysokých škol ve světových bázích dat byla vyvolána především zvýšeným zájmem o hodnocení jednotlivých škol a případně jejich složek nezávisle na údajích, které poskytují samotné školy. Ukazuje se, že tento problém nelze oddělovat od publikační aktivity jako takové, ovšem ve světovém měřítku, protože informační zdroje s takovou šíří pracují a snaží se, aby skutečně byla co největší. Přesto, že jsme se věnovali s pochopitelných důvodů jen vlastní vysoké škole, kde máme jak oprávnění, tak i možnost individuální kontroly nalezených dat, a že i při jistém zobecnění se budou naše výsledky vždy spíše týkat chemie jako oboru s přijatelným rozšířením na přírodní vědy, lze konstatovat, že výsledky z obou bází jsou v zásadě totožné, neboli jinými slovy, obě báze dat pokrývají celosvětovou produkci vědeckých publikací prakticky ve stejné šíři. Rozdíly jsou celkem snadno vysvětlitelné a vyplývají z mírně rozdílných výchozích koncepčních zásadách

a je tudíž snadno možné se rozhodovat, zda bude v daném případě dáována přednost té či oné koncepci. V neposlední řadě samozřejmě hraje roli i nástroj, jehož prostřednictvím uživatel s danou bází dat pracuje, kde ale hrají roli spíše osobní preference a také samozřejmě i zvyk na tu či onu bází dat, resp. příslušný vyhledávací nástroj nebo rozhraní. Budiž tedy konstatováno, že pokud budeme chtít posuzovat publikační aktivitu univerzit a podobných institucí z hlediska světové vědy, je pokrytí naší produkce v zásadě odpovídající a prakticky plně srovnatelné necht' je pro získání dat využita báze dat Web of Science nebo Scopus.

Reference:

1. BERGER, J.: DNES, 6.června 2007.
2. BERGER, J.: DNES, 14. listopadu 2007.
3. ANON.: iDNES, 20.června 2007
4. ŠILHÁNEK, Jaroslav: Orientační testování báze dat Scopus. Ikaros, 11, č. 1 (2007), <http://www.ikaros.cz/node/3800>.
5. LaGUARDIA, Cheryl. 2005. E-Views and Reviews: Scopus vs. Web of Science. *Library Journal* [online]. Jan. 15, 2005 [cit. 2006-12-29]. Dostupné na World Wide Web: <<http://www.libraryjournal.com/article/CA491154.html%22>>
6. Během přípravy tohoto článku oznámilo nakladatelství Elsevier další nástroj pro práci s bází dat Scopus, a sice tzv. „Scopus Affiliation Identifier“, který umožňuje automaticky identifikovat alternativní adresy.
7. Scopus overview: <http://www.info.scopus.com/>