

VOLNÉ SDÍLENÍ ODBORNÝCH ČLÁNKŮ – SCI-HUB PIRÁTSTVÍ NEBO TREND?

Jiří Kofránek, Richard Papík

Abstrakt

Informační technologie zásadně mění dosavadní způsob publikování odborných sdělení i vydavatelské obchodní modely. Prudce se rozšiřuje podíl open-source časopisů. Zároveň vznikají komunity, jejíž členové si vzájemně vyměňují publikované články a obcházejí tak placené časopisecké přístupy. Vznikají také i pirátské komunity, které zásadním způsobem narušují autorská práva vydavatelů – jako např. Sci-Hub, těžko se ale tomu dá bez přispění tvrdých restriktivních opatření garantovanými státem zamezit. Ve vědecké komunitě tak vzniká obdobná situace, jako i v minulosti s pirátským kopírováním písní a filmů. Řešením v budoucnu nebude tvrdá restrikce, ale změna vydavatelské obchodní politiky.

Klíčová slova

internet, časopisy, impact factor, open source, Sci-Hub, pirátství

1 Current Contents a noční kopírování

Když jsem jako mladý doktorand koncem sedmdesátých let přišel na Ústav patologické fyziologie, každotýdně kolovala mezi pracovníky mnohastránková brožurka tištěná na tenkém cigaretovém papíře s názvem Current Contents. Byl to časopis vydávaný Institutem vědeckých informací (Institute for Scientific Information - ISI) ve Philadelphii. Brožura obsahovala seznamy publikovaných článků vědeckých časopisů, které ten daný týden vyšly spolu s adresami autorů, kterým bylo možné poslat žádanku o poskytnutí kopie (tzv. separátního výtisku) daného článku. A tak jsme každý týden vyplňovali do předtištěných korespondenčních lístků název článku a adresy autorů. A po čase kupodivu přišly obálky s vyžádanými články jako odpověď na rozeslané korespondenční lístky. Existovala ještě další možnost, pokud potřebný časopis odebírala knihovna – podle údajů získaných z brožurky na cigaretovém papíře bylo možné si článek přečíst v knihovně, a dokonce si v knihovně objednat i jeho kopii. S kopírkami ale byl tehdy problém, protože se tehdejší režim bál jak čert kříže šíření nepovolovaného obsahu – kopii článku mohla udělat jenom knihovna na žádanku. Koncem devadesátých let, kdy byla v prostorách Ústavu patologické fyziologie usídlena Hybridní výpočetní laboratoř, součástí výpočetního střediska byla i kopírka. Tehdejší šéf této laboratoře, ing. Kotva, umožnil samostatné kopírování na této kopírce všem zaměstnancům lékařské fakulty, což byl na tehdejší dobu malý zázrak. Mohli jsme si tehdy půjčit večer v Národní lékařské knihovně poslední číslo odborného časopisu (který byl jen pro prezenční čtení) a do rána si na kopírce vlastníma rukama zkopírovat potřebný článek.

2 Vědeckocitační rejstříky

Za vznikem oné mýtické brožurky, z níž jsme kdysi čerpali informace o nových odborných článcích, stál legendární Eugene Garfield. Původně studoval na Columbia University chemii, z níž obdržel bakalářský titul, magisterský titul získal v roce 1954 v knihovnické vědě. Doktorát obdržel v roce 1961 v oboru lingvistiky na University of Pennsylvania. Od roku 1955 pracoval jako konzultant pro farmaceutické firmy a brožurka Current Contents, kterou v té době založil, se původně věnovala farmacii a příbuzným oborům. V roce 1955 v článku v časopise Science [1] přišel s ideou citačního rejstříku – databáze, v níž by se systematicky shromažďovaly provázané citace vědeckých publikací a která by umožnila odlišit nejlepší vědecké časopisy od těch ostatních. Navrhl v podstatě jednoduchý systém „označkování“

úrovně časopisů pomocí impaktních faktorů (Impact Factor), založený na porovnání průměrné citovanosti článků uveřejněných v daném časopise. Navrhl (i prakticky realizoval) i metodiku vytváření a ukládání těchto informací, založenou na děroštitkově technologii IBM [2]. V roce 1960 založil výše zmíněný Institut pro vědecké informace (ISI) a v tomto institutu svou myšlenku vědeckocitačního rejstříku a impaktních faktorů prakticky realizoval – v roce 1963 vydal svůj první Science Citation Index, který umožnil pomocí impakt faktoru měřit význam odborných časopisů a sledovat citační ohlas prací jednotlivých autorů. Po prvních tištěných experimentálních číslech v roce 1963 byl SCI vydáván pravidelně: od roku 1964 nejprve 4krát a pak již 6krát ročně s roční kumulací na konci roku. Garfieldův institut ISI vydával již výše zmíněný každotýdní časopis Current Contents pro řadu oblastí vědy i další citační rejstříky: vedle Science Citation Index (SCI) od roku 1973 také Social Sciences Citation Index, Arts & Humanities Citation Index (od roku 1978) a Index Chemicus. Garfieldova práce vedla k vývoji několika algoritmů pro hodnocení důležitosti informací jako např. HITS a Pagerank.

„SCI umožnil svým pojetím představit výzkum jako formalizovanou síť propojenou cestou slov a citací, čímž se splnil Garfieldův sen. Bylo jím vytvořit kombinaci konceptu SCI, počítačů a literatury zcela nový svět, založený na semiotice citování. Garfield, původním vzděláním chemik, ve své snaze po získání kamene mudrců přeměnil olovo (reference) ve zlato (citace). Přitom on sám není ani chemik obrácený na dokumentátora, není ani informační vědec nebo IT vynálezce.

Představuje bohaté a tvořivé tradice stejně tak staré jako jsou citační rejstříky. E. Garfield je Alchymista“ – psal o Garfieldovi v roce 2006 Milan Špála [3], který se s Garfieldem osobně znal a zasloužil se uplatnění jeho myšlenek v československém a českém kontextu. Garfield několikrát navštívil Prahu, a v roce 1995 obdržel čestný doktorát Univerzity Karlovy.

V roce 1992 Gafield odešel do důchodu a jeho institut se pod značkou Thomson ISI stal součástí společnosti Thomson Reuters, která začala provozovat on-line verze vědeckocitačních rejstříků pod názvem Web of Science. V roce 2016 byl Thomson ISI prodán společnosti Clarivate Analytics. Koncem roku 2004 se narušil do té doby monopolní zdroj citačních ohlasů Web of Science, kdy nizozemská firma Elsevier zavedla novou databázi Scopus, která rovněž eviduje citační ohlasy.



Obrazek 1 – Eugene Garfield (1925-2017), fotografie z roku 2007.

Garfieldovy rejstříky a návazné databáze kvantifikující úroveň časopisů podle impaktního faktoru zásadně ovlivnily trh odborných a vědeckých časopisů. Impaktní faktor se stanovuje ze dvou položek: čitatele a jmenovatele, číteců a počet citacích ohlasů v běžném roce k jakémukoliv z článků, které daný časopis (patřící do sledované databáze časopisů) celkem uveřejnil v předchozích dvou letech (nyní i v předchozích pěti letech). Jmenovatel uvádí celkový počet článků za ono dvouletí (pětiletí). Pro každý vědecký časopis je proto důležité, zda je, či není v databázi se sledovaným a přiděleným impaktním faktorem. Impaktovaným časopisům pak obvykle špičkoví autoři dávají přednost, a tudíž si editoři mohou vybírat z většího množství příspěvků, což se pak může odrážet v kvalitě publikovaných článků.

Přiosvětila se tím konkurence - trh vědeckých a odborných časopisů se rozdělil na časopisy „impaktované“ a na ty ostatní, „neimpaktované“. Zařazení mezi impaktované časopisy však není snadné (přes 2 tisíce žádostí ročně, úspěšnost cca 10 %), zvláště pokud mezi impaktovanými časopisy již existuje časopis se stejným tematickým zaměřením [4].

Nové informační technologie, které poslaly do muzeí sázecí stroje a umožnily elektronickou sazbu a zjednodušily tisk, konkurenci ještě dále zostřily.

K tomu přispěla i kritéria grantových agentur a státních institucí zvláště v Evropě hodnotící úroveň vědeckých pracovníků podle toho, jaký impaktní faktor má časopis, ve kterém publikovali své články.

Jak poznamenal Milan Špála, Garfield ve své práci o historii a významu časopiseckého impaktního faktoru tvrdil, že v roce 1955, kdy impaktní faktor navrhl, netušil, že by se mohl stát v budoucnu tak kontroverzním. Předpokládal, že bude používán konstruktivně, a nemohl tušit, že ve špatných rukou může být zneužit [3]. V roce 1998 E. Garfield považoval za nutné začít přednášku na výročním shromáždění významné instituce, jakou je „Council of Biology Editors“ (CBE), větou: „*Používání impaktního faktoru, zvláště v Evropě, vzbudilo mnoho obav*“ [5]. A uvedl, že je nebezpečné aplikovat jakékoli statistické údaje mimo kontext (impaktní faktor (IF) je klasickým výsledkem statistických operací); IF by se neměl používat jako náhražka skutečné citační výkonnosti (vysoký IF časopisu, v němž někdo publikoval článek nezaručuje, že tento článek získá skutečně vysoký počet citací, protože IF je statistickým průměrem všech citací daného časopisu, a trpí rizikem zkreslení díky ojedinělým citačně úspěšným článkům, jejichž zasloužený obdiv se použitím „průměrného“ IF, a ne skutečných citací, neoprávněně přenesl i na článek, citovaný třeba hluboce pod průměrem).

Garfield v den svých osmdesátých narozenin, 16. září. 2005, v Chicagu na Mezinárodním kongresu o peer review a lékařském publikování ve své přednášce, již se kongres otevíral, považoval za nutné při uvažování o různých aplikacích znovu upozornit na zásadní kontroverzalitu nahrazování počtu skutečných citací jen IF časopisu. V této přednášce, opublikované o rok později časopisecky [6], doslova píše: „*Grantové agentury a další instituce si často přejí vyhnout se práci se zjišťováním počtu citací na jednotlivé články a autory. Čerstvě zveřejněné články údajně nemají dost času být citovány, a tak se pokoušejí použít i. f. jako náhradní evaluační nástroj. Tím se předstírá, že již samotné přijetí článku k uveřejnění v časopise s vysokým i. f. je oním předpokládaným indikátorem věhlasu. A tak se i. f. používá k odhadu předpokládaného významu jednotlivých článků, který je naopak spíše pochybný, vzhledem ke známým odchylkám i. f. většiny časopisů.*“

3 Internet mění trh odborných a vědeckých časopisů

Prudký rozvoj Internetu vedl k tomu, že čtenář vlastně nepotřebuje tištěnou verzi časopisu, a stačí mu elektronická verze

článku. Požadavkem doby se stalo mít (krom tištěné verze) i elektronickou verzi časopisu.

Velkým záskem do trhu vědeckých časopisů byl rozvoj internetových vyhledávačů – zejména Googlu. V listopadu 2004 Google přichází s beta verzí vyhledávače Google Scholar, který se velmi rychle zabydlil v akademické komunitě. Vyhledávací a třídící algoritmy Google Scholar se rok od roku zlepšovaly. Google Scholar dnes zahrnuje citační odkazy na většinu recenzovaných akademických časopisů, citační odkazy na knihy, konferenční příspěvky, diplomové práce a disertace, preprinty, abstrakta, technické zprávy, a jinou odbornou literaturu, včetně patentů. Pokud je veřejně k dispozici elektronická verze plného textu, poskytne na ní odkaz.

Snadnost vyhledávání a šíření elektronickou cestou zásadně změnila vydavatelské zvyklosti. Každý renomovaný časopis má elektronickou verzi, za jejíž zpřístupnění ale většinou požaduje určitou platbu (v rozmezí cca 30-40\$ za článek). Aby se pokryly náklady, stoupají ceny předplatného. Stoupají i publikační poplatky, které platí autor za publikaci přijatého článku do renomovaného časopisu.

Objevily se odborné časopisy, které nabízejí texty odborných článků zdarma. Aby se pokryly obchodní náklady, jsou poměrně vysoké publikační poplatky, které platí autor za přijatý článek. Brzy se ukázalo, že otevřená dostupnost zvyšuje citovanost článku (samozřejmě za předpokladu, že editor dbá na vysokou úroveň recenzního řízení). Renomované tištěné časopisy nabízejí otevřené zpřístupnění článku za cenu zvýšení vstupních autorských poplatků.

Příkladem poměrně dravého uvedení na trh otevřeného časopisu je nakladatelství JMIR Publication – v roce 1999 toto nakladatelství uvedlo na trh recenzovaný časopis **Journal of Medical Internet Research** (<http://www.jmir.org/>) věnovaný problematice eHealth v internetové éře. Rychlost (ale i kvalita) recenzního řízení a publikace výlučně na internetu vedli k tomu, že články začaly být citovány a časopis několik let sledoval jejich citovanost a urputně uváděl „neoficiální Impact Factor“, který stále stoupal, až po cca sedmi letech byl přijat mezi oficiálně sledované „impaktované“ časopisy – dnes s IF 5.175 patří k lídrům časopisů věnovaným digitální medicíně a elektronickému zdravotnictví. Obchodní model je založen na poměrně vysokých vstupních poplatcích (a významných slevách těchto poplatků pro recenzenty).

Jiným příkladem úspěšného otevřeného časopisu je **PLOS one** (<http://journals.plos.org/plosone/>). Základem jeho úspěchu je kvalitní recenzní řízení a doprovodné elektronické služby. Dalším příkladem nových otevřených časopisů jsou časopisy z nakladatelství **Frontiers** (<https://www.frontiersin.org/>). Základem úspěchu je náročné recenzní řízení (s otevřenou publikací recenzentů, takže kvalitní recenze, která přispívá ke zlepšení článku je dohledatelná), možnost dodatečného „supplementary“ elektronického materiálu, možnost veřejné elektronické diskuze aj.

Na druhé straně ovšem relativní snadnost publikování vedla také k rozvoji tzv. „predátorských časopisů“, které vznikají primárně s cílem vybírat autorské publikační poplatky a generovat zisk, nikoli podporovat a rozvíjet vědeckou komunikaci. Vznikají i seznamy „predátorských časopisů“ – viz např. <https://beallslist.weebly.com/>, publikací, kterým by se měli akademici vyhybat. Někdy je ale zařazení do obdobných „černých seznamů“ sporné a vyvolává velké polemiky až soudní spory. V nedávné době vznikla řada nových časopisů, zejména v Číně a dělat rigorózní soudy o jejich úrovni bez podrobného rozboru může být zavádějící – jejich odborná úroveň se ukáže časem.

4 Dostupnost elektronického obsahu časopisů

Ceny předplatného vědeckých časopisů stoupají, a řada univer-

zit (zejména v Evropě) začíná šetřit. Ačkoli je mnoho informací veřejně dostupných přes internet, tak v případě odborných článků to příliš neplatí. Pouze zhruba méně než desetina publikovaných článků je po svém zveřejnění okamžitě dostupná přes otevřený přístup (Open Access) – stačí se podívat na Google Scholar. A platit třicet dolarů za článek, který není v předplatném vědecké instituce, je hodně.

Nejprve vědecká instituce vynaloží veřejné prostředky na výzkum, publikuje výsledky a pak ale za zpřístupnění vědeckých výsledků (které jiná instituce rovněž vytvořila za veřejné prostředky) musí platit. Před dvěma lety se v rámci Evropské unie sešli příslušní ministři jednotlivých zemí a shodli se, že by bylo velmi potřebné z hlediska rozvoje Evropské unie, aby tato bariéra padla [7]. Ale představa, že od roku 2020 všechny publikované články, jejichž výzkum je placen z veřejných finančních prostředků, musí být také veřejně dostupný, je jen deklarací. A navíc řada vydavatelů prestižních vědeckých publikací má sídlo v USA.

A v USA taková benevolence není – vzpomeňme na Aarona Schwartze, který byl v roce 2011 obviněn z toho, že přes otevřené připojení v Massachusettském technologickém institutu (MIT) údajně ukradl více než čtyři miliony akademických časopisů z digitální knihovny JSTOR a hodlal je veřejnosti poskytnout zdarma. Byl obžalován v celkem 13 bodech a hrozilo mu 30 let ve vězení a pokuta ve výši milionu dolarů. Žalobu proti Swartzovi iniciovalo ministerstvo spravedlnosti Spojených států. Ačkoli JSTOR žalobu stáhl poté, co slíbil, že získané informace nebude šířit dál, proces pokračoval a stále mu hrozilo až třicet let vězení. Aaron Schwartz 12. ledna 2013 ve svých 26 letech spáchal sebevraždu oběšením (https://cs.wikipedia.org/wiki/Aaron_Swartz).

Rozvoj sociálních sítí vedl k vzniku skupin vědců, kteří si v těchto sítích vyměňují své články. Příkladem je **Academia** (<https://www.academia.edu/>) nebo **ResearchGate** (<https://www.researchgate.net>). Výměna článků je zatím tiše tolerována vydavateli, i když se často jedná o porušení licenční smlouvy, v níž se nejčastěji uvádí, že lze veřejně vystavit pouze první verzi článku, která byla do časopisu zaslána. Avšak dle licenční smlouvy, kterou autor podepsal, nelze publikovat poslední verzi, která byla upravena dle (několika kol) recenzních posudků. Když by chtěl autor volně se svým článkem nakládat, tak by měl zaplatit poplatek (článek je pak ve volném režimu Gold OpenAccess), jehož výše je často hodně vysoká, až kolem cca 2000€. Autoři ale toto

licenční ujednání často porušují a zveřejňují licencované články ve veřejných repozitářích nebo na akademických sociálních sítích.

Tolerance ale není trvalá. Skupina pěti velkých vydavatelů před časem varovala z Německa pocházející síť ResearchGate, že sdílení prací považuje za porušení vlastnických práv. Německá služba pak v tichosti začala takové práce odstraňovat (resp. blokovat k nim přístup), což „Koalice pro zodpovědné sdílení“, jak se zájmová organizace vydavatelů nazývá, na začátku října 2017 kvitovala s uspokojením. O týden později své výzvy sice zopakovala, ale případ nadále neeskaloval a obě strany zatím prý vedou „vážné rozhovory“. ResearchGate má zhruba 13 milionů členů, kteří na jejich stránky umístili desítky milionů prací.

5 Pod pirátskou vlajkou

Roku 2011 se to rozhodla změnit Alexandra Elbakyanová z Kazachstánu; tyto ceny jí připadaly nestydaté a s argumentem, že „věda by měla být dostupná všem“ stvořila Sci-Hub. Elbakyanová je studentka neurověd a programátorka, s pomocí přátel proto vytvořila repozitář vědeckých akademických článků, kde se shromažďuje velké množství studií z mnoha prestižních vědeckých magazínů.

Funguje to tak, že anonymní akademici z celého světa poskytují Sci-Hubu své předplatné, díky němuž server získává články. Ty jsou pak na něm zcela zdarma přístupné úplně všem. Nyní (březen 2017) je na Sci-Hubu přístupných 64,500,000 článků a jejich počet stále roste. Je jasné, že služba Sci-hub, která nabízí „upíratěně“ vědecké studie zdarma ke stažení, čelí mnoha žalobám.

Tak např. v roce 2015 žalovala Sci-Hub společnost Elsevier a v červnu 2017 byl Sci-Hub odsouzen k pokutě 15 miliónů dolarů za porušování autorských práv. Krom toho, v červnu 2017 soud v americké Virginii souhlasil s žalobou Americké chemické společnosti (America Chemical Society – ACS) a odsoudil službu Sci-Hub za porušování autorských práv k pokutě 4,8 miliónu dolarů. Rozsudek požaduje, aby se Sci-Hub přestaly spolupracovat všechny organizace či služby, které by například mohly na svých serverech přechovávat ukradené práce. Teoreticky by se opatření mohlo vztahovat i na vyhledávače, které návštěvníky na Sci-Hub zavedou, ale v tomto ohledu není podle všeho právní názor jednotný. Dá se těžko čekat, že vyhledávače by takový



Obrázek 2 – Úvodní stránka pirátského serveru Sci-Hub



Obrázek 3 – Zakladatelka serveru Sci-Hu, Alexandra Elbakjanová na konferenci na Harvardově univerzitě v roce 2010, foto: Apneet Jolly, CC-BY, Creative Commons.

krok dobrovolně udělaly, protože by to pro ně byl velmi nepřijemný precedent.

Pirátská služba Sci-Hub mnohokrát změnila svou adresu – nyní je hostována na adrese Taiwanské domény <http://sci-hub.tw/>. Krom toho je zatím ještě dostupná na adresách: <http://sci-hub.hk/>, <http://sci-hub.name/>, <http://sci-hub.tv/>, <https://sci-hub.la/>.

Je otázkou jak dlouho bude dostupná – v každém případě přináší velkou škodu tradičním vydavatelům jako je Elsevier a další, nicméně ukazuje cestu do budoucna.

6 Jak dál

Boj velkých vydavatelů proti pirátskému Sci-Hubu může být úspěšný v USA a při velkém nátlaku vlastníků velkých vydavatelství i v Evropě. Těžko lze asi zcela potlačit sdílení vědeckých článků, pokud by se nezavedla přísná regulace internetu. Volné sdílení odborných publikací není nějaký dočasný exces typu Sci-Hubu, ale trend.

Je zřejmé, že současný model dostupnosti vědecké literatury je z dlouhodobého hlediska neudržitelný. Tradiční obchodní modely tradičních vědeckých časopisů – předplatné pro knihovny a poplatky stažení článků budou muset skončit. Podíl volně stažitelných časopisů bude stoupat, a budou se hledat nové cesty financování. Cesta bude možná i v zapojení cílené reklamy (obdobně jako Google, nebo Facebook) do sociálních sítí pro sdílení publikací nebo přímo do stahovaných elektronických publikací – jakou cestou se půjde, to ukáže budoucnost.

Literatura

- [1.] Garfield E. *Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas*. Science. American Association for the Advancement of Science; 1955;122: 108–111.
- [2.] Garfield E. *The preparation of printed indexes by automatic punched-card techniques*. Journal of the Association for Information Science and Technology. Wiley Online Library; 1955;6: 68–76.
- [3.] Špála M. *Impakt faktor - Dobrý sluha, ale špatný pán*. Časopis lékařů českých. 2006;145: 69–78.

[4.] Pravda V, Křížek M. *Citace: dobrý sluha, špatný pán*. Pokroky matematiky, fyziky a astronomie. Jednota českých matematiků a fyziků; 2007;52: 28–36.

[5.] Garfield E. *The use of journal impact factors and citation analysis for evaluation of science*. 41st Annual Meeting of the Council of Biology Editors, Salt Lake City, UT. [garfield.library.upenn.edu](http://www.garfield.library.upenn.edu/papers/eval_of_science_oslo.html); 1998. Available: http://www.garfield.library.upenn.edu/papers/eval_of_science_oslo.html

[6.] Garfield E. *The history and meaning of the journal impact factor*. JAMA. jamanetwork.com; 2006;295: 90–93.

[7.] *All European scientific articles to be freely accessible by 2020*. In: NL EU [Internet]. 2016. Available: <https://www.politico.eu/wp-content/uploads/2016/05/NLopenaccess.pdf>

Kontakt

Jiří Kofránek

Oddělení biokybernetiky a počítačové podpory výuky
ÚPF 1. LF UK
U Nemocnice 5 128 53
Praha 2
email: kofranek@gmail.com