

Stínoví knihovníci

Sen o informačním universu

Pirátská komunita je tvořivá. Nejenže získává a šíří informace, ale také je přepracovává, katalogizuje, spravuje a třídí. V posledních letech s piráty našli společnou řeč i vědci a všichni, kdo se potřebují dostat k odborným publikacím.

HONZA ŠÍPEK

Web, tedy celosvětová síť, se zrodil v CERNu v roce 1990 ze snu o informačním universu, které by vědcům poskytovalo přístup k odborným poznatkům. Systém měl umožňovat například postup od citací k citovanému zdroji. „Představte si, že se dotknete odkazu v článku,“ píše zakladatel webu Tim Berners-Lee dva roky po svém objevu. „Předpokládejme, že se vám okamžitě objeví informace z jeho pozadí – další články, údaje o autorovi, adresa jeho instituce i celý její telefonní seznam.“ Web brzy protkal celý svět a stal se důležitým společenským i komerčním hybatelem. Poznanky vědeckého výzkumu jsou v něm ale schované za „platební zdí“ (paywall), a ta někdy bývá pěkně vysoká.

Vědci jsou lapani v podivném paradoxu. Jejich výkon je totiž posuzován za pomoci „scientometrie“ – počtu zveřejněných odborných článků a jejich citací. Citační index je něco jako skóre v počítačové hře, které otevírá další levely kariéry. Má to svůj smysl: příspěvky do časopisů procházejí často náročným recenzním řízením a počet citací může napovídát něco o tom, jak důležitý byl výzkumníkův objev pro zbytek světa. Většinu odborných článků ale vydává pět největších nakladatelství: Reed Elsevier, Springer Nature, Wiley-Blackwell, Taylor & Francis a Sage. Honoráře dávno neplatí ani autorům, ani recenzentům a většinou chtějí plná práva, takže vědec sám nemůže text zveřejnit. Články se pak publikují především v elektronické podobě, kterou ale nabízejí ke čtení za vysoký poplatek. Vědecké ústavy, univerzity a knihovny si proto předplácí pouze několik časopisů, které si mohou dovolit (jde o desetitisíce až statisíce korun ročně). Dochází tak k paradoxním situacím. Například nedávno publikovaný článek laserového fyzika zaplatil daňový poplatník několikrát: jednou, aby vznikl, podruhé, když si ho čte některý z vědců, a potřetí, když si článek chce přečíst jeho kolega z jiného ústavu.

Náklady na knihovnické služby se tak stávají čím dál tíživější položkou rozpočtů vědeckých ústavů. Poměrně dobře jsou na tom vědci a studenti z rozvinutých zemí – k nějakým časopisům se dostanou. Co ale vědci z třetího světa, kteří mají jen na základní pomůcky a všechno ostatní je pro ně drahé? Nemluvě ani o tom, jak se ke článku dostane laik, učitel, novinář či prostě jen zvědavý občan, z jehož daní se výzkum platil. Za každý článek musí zaplatit cenu od třiceti do čtyřiceti dolarů. Za text v rozsahu přinejlepším desítek stran, který jeho autor odevzdal redakci zadarmo, se tedy platí stejně jako za výpravnou publikaci na křídovém papíře.

Mučedník internetu

Aaron Swartz ukradl oheň bohům a stal se prvním mučedníkem internetu. Tento talentovaný hacker stál u zrodu systémů jako Reddit, RSS či tor2web. Veden přesvědčením, že informace mají patřit všem, už dříve „scrapoval“ veřejné dokumenty amerických federálních soudů z veřejné databáze PACER. Jím vytvořený skript, scraper, vyhledával a stahoval soudní dokumenty pečlivě jeden za druhým, ale mnohonásobně rychleji a efektivněji, než by to dokázal člověk rychlým klikáním. Swartz byl za stažení 2,7 milionu dokumentů

vyšetřován, ale obvinění byla stažena, protože rozhodnutí soudu nejsou chráněna copy-rightem, a co je přístupné všem, nelze ukrást.

Na podzim 2010 přednášel Swartz v Budapešti na konferenci o svobodě internetu. Potkal se tam s aktivisty, kteří chtěli za peníze získané od mecenáše George Sorose zpřístupnit veřejnosti celou placenou databázi odborných časopisů JSTOR. Zjistili ale, že zaplacení poplatků by vyšlo na stovky milionů dolarů. Swartzovi se v hlavě začal rodit plán.

Bydlel nedaleko kampusu americké technické školy MIT, která je mimo jiné kolébkou hackerů. Tato bohatá škola měla přístup ke kompletní databázi JSTORu. Swartz sice na MIT nestudoval, ale návštěvníci kampusu měli volný přístup k počítačové síti a oficiálně i ke knihovnickým službám. První pokusy se scraperem JSTOR odřízl, protože přetěžovaly jeho servery a robotické stahování bylo proti uživatelským podmínkám. Swartz tedy scraper upravil tak, aby stahoval články opatrněji. Ve volně přístupném sklepě budovy 16 vlezl do nezamčeného kumbálu se síťovými rozbočovači. Připojil notebook do switche, schoval ho na polici pod papírovou krabici a nechal stroj pracovat. Program postupně stahoval miliony článků. Stačilo jednou za čas přijít vyměnit disky. JSTOR si ale stahování všiml a dal podnět k vyšetřování. Do hry se vložila i tajná služba a skrytá kamera schovaná na místě natočila Swartze právě při výměně disků. Po krátké honičce byl dopaden.

Protože návštěvníci kampusu měli volný přístup k článkům a hacker porušoval vlastně jen uživatelské podmínky JSTORu, byl nakonec obviněn z neoprávněného přístupu k síti. Pod tlakem dlouhého vyšetřování se Swartz 11. ledna 2011 ve svých šestadvaceti oběsil a internet získal svého prvního mučedníka.

Stínové knihovny

Knížky se tehdy sdílely podobně jako ostatní „warez“, tedy pirátský software, muzika a filmy. Do anonymního fóra, přístupného pouze pod heslem, poslal někdo odkaz na soubor na stahovacím serveru. Způsob je to dost nepřehledný – soubory se objevují chaoticky z různých oborů a od různých autorů. Linky

na stahovacích serverech jsou plné reklamy a po nějaké době vyprší a zmizí. „Prosím re-up!“ žadoní frustrovaní uživatelé. Taková situace přímo vybízí k vytváření domácích, poloveřejných i zcela otevřených knihoven. Některé začaly být tak velké, že přitáhly tisíce nadšených uživatelů, ale také pozornost „orgánů“. Rok po smrti Aarona Swartze vyhověl mnichovský soud stížnosti sedmnácti velkých nakladatelů a nařídil uzavřít webovou knihovnu Library.nu. V době svého uzavření obsahovala asi 400 tisíc knih, což odpovídá Alexandrijské knihovně.

Akademici a výzkumníci z celého světa zaplakali: byly tam učebnice, monografie, biografické analýzy, manuály, výsledky výzkumu v technice, matematice, biologii, sociálních a humanitních vědách. Dokonce i pornografie tam byla vědecká: průvodce a učené knihy o pornoprůmyslu. „A to, že se podzemní zločinecká stránka obejde bez pornografie, musíme považovat za jeden z triumfů civilizace,“ napsal profesor Christopher Kelty z UCLA.

Sen o univerzální knihovně se mezitím zhmotňoval v projektu s názvem Library Genesis. Nápis v azbuce na stránce libgen.org dávaly tušit provenienci serveru. Knihovna nejenže obsahuje miliony odborných knih, ale i jejich bibliografické záznamy s dobrým vyhledáváním. Všechny soubory jsou přístupné přes torrent, takže stažené kopie mohou od uživatele putovat dál i v případě zmizení původního zdroje. A server jde v otevřenosti ještě dál: vystavuje i „dump databáze“, ze kterých může svou „knihovnu stvoření“ sestavit každý kdekoli jinde. Tedy každý, kdo si pořídí čtyřicetiterabajtové diskové pole. Klony knihovny již vznikly na různých adresách. Nekonkurují si, nýbrž spolupracují, takže libgen.org nabízí soubory ke stažení z více možných „zrcadel“. Hodí se to zvláště při častých výpadcích hlavního serveru.

Bitva o Sci-hub

Alexandra Elbakjan, studentka neurovědy z Kazachstánu, se vydala na univerzity do německého Freiburgu a do Spojených států. Zvykla si na tamní vědu i přístup k informacím, jenže po návratu do vlasti byla odříznutá

od svých známých informačních zdrojů. Když její kolegové potřebovali nějaký odborný článek, poprosili o něj v polotajných diskusních fórech, a pokud měli štěstí, kolega z instituce se zaplaceným přístupem k časopisu jim ho stáhl a poslal. Na Twitteru začal podobně fungovat tag #ICANHAZPDF. Elbakjan jako šikovná hackerka přemýšlela, zda by nešel tento systém zautomatizovat, a naprogramovala Sci-hub. „Nebyla za tím žádná velká myšlenka jako ‚osvobodte informace‘. K výzkumu jsme prostě jenom potřebovali číst články,“ říká v rozhovoru pro server TorrentFreak. Do vyhledávací stránky napíšete jméno článku. Pokud ho Sci-hub již má v databázi, nechá vás ho stáhnout. Pokud ne, pošle impuls do pavučiny tajných robotů kdesi v pozadí. Ty mají přístup do různých databází nakladatelů za platební zdí, článek vám stáhnou a schovalí kopii pro ty, kteří ho budou hledat příště. Přístupové údaje prý věnovali vědci – buď dobrovolně, nebo nevědomě.

Soudy sice již nařídily zabavení domén libgen.org i sci-hub.org, ale americká justice má moc pouze nad americkými doménami, a tak se brzy oba weby objevily na nových adresách. Když byly zakázány i ty, pokračovaly jinam. Takovou doménovou hru na kočku a myš hrají soudy v The Pirate Bay už celé roky, a Pirátská zátoka pořád existuje.

Dnes osmadvacetiletá Elbakjan se nijak zvláště neskrývá – tají pouze místo pobytu, ale má se za to, že žije „někde v Rusku“, mimo dosah amerických soudů. Dala rozhovor i prestižnímu časopisu Science a po anonymizaci mu ochotně předala i data o používání svého webu. Z analýzy časopisu vyšlo najevo, že kromě třetího světa, který hledá cestu k informacím z medicíny a technologií, používají web ve velkém i západní vědci a studenti. Je to jednodušší, zadarmo a najdou to, co potřebují, bez ohledu na škrty v rozpočtech ústavní knihovny. Buď se s tím nakladatelé vyrovnají tak, jako se hudební průmysl vyrovnal s Napsterem, a zásadně změní svou strategii, anebo nad nimi vyhraje sdílená stínová knihovna, která je tím robustnější, čím víc se jí snaží eliminovat.

Autor je dokumentarista.



Aaron Swartz na záběru z videa U. S. Secret Service, které ho mělo usvědčit