

# Čo nám povedia čísla alebo knižničná činnosť z pohľadu štatistik

Jana ILAVSKÁ

Univerzita Komenského v Bratislave - Akademická knižnica, Bratislava

[jana.ilavska@rec.uniba.sk](mailto:jana.ilavska@rec.uniba.sk)

INFORUM 2011: 17. konferencie o profesionálnych informačných zdrojích  
Praha, 24. - 26. 5. 2011

## Abstrakt

*V rámci hodnotenia činnosti knižníc sa často využívajú rôzne štatistické ukazovatele. Činnosť slovenských akademických knižníc je posudzovaná jednak na základe oficiálnych štatistických výkazov Ministerstva školstva, vedy, výskumu a športu SR ako aj na základe rôznych prehľadov a tabuliek, ktoré sú súčasťou výročných správ univerzít. Viaceré vysoké školy žiadajú od svojich knižníc ďalšie podklady za účelom vnútorného hodnotenia pracovísk, fakúlt, katedier, jednotlivcov. Príspevok prinesie prehľad rôznych štatistik, ktoré vytvára Akademická knižnica Univerzity Komenského v Bratislave, upozorní na úskalia získavania vstupných informácií a na problémy pri ich vyhodnocovaní. Interpretácia tých istých výsledkov nemusí byť vždy rovnaká. Mnohokrát závisí od uhla pohľadu a od účelu, pre ktorý sa údaje žiadajú.*

Keď sa stretne s termínom štatistika, väčšina si predstaví nudné čísla. Niektorým sa možno v mysli vybaví pomerne bežné kliše, že štatistika je superlatívom lži. Avšak nemožno poprieť, že štatistika je veľmi dôležitá, a je to tak aj v knižnično-informačnej oblasti. Štatistické ukazovatele často používame pri obhajovaní opodstatnenia činnosti, dokonca niekedy aj pri obhajobe existencie knižnice. Je len na nás, či ich podáme enumeratívne alebo spôsobom, ktorý pritiahne pozornosť. Aj čísla môžu rozprávať príbehy. A čo sa nedá zmerať, či zrátať, sa často nepočíta. Z rôznych druhov štatistik, ktoré knižnica vyhotovuje, sa sústreďíme najmä na štatistiky odzrkadľujúce hodnotu, či prínos knižnice pre komunitu, v ktorej pôsobí. Sú to najmä výkonové ukazovatele a využívanie elektronických informačných zdrojov.

Pred prezentáciou a interpretáciou údajov, musí prebehnúť proces ich zberu a spracovania. Vo všeobecnosti je potrebné rozlišovať medzi termínmi dáta/údaje – informácie – štatistika. Sú to odlišné pojmy, ktoré sa však často chápu nesprávne. Dáta môžeme charakterizovať ako „surovinu“ na výrobu informácií. Sú to údaje, fakty alebo čísla, z ktorých možno vyvodit' závery. Za informácie môžeme považovať údaje, ktoré boli zaznamenané, klasifikované, organizované, dané do súvislostí alebo interpretované v nejakom rámci, z ktorého vyplynul ich zmysel a význam. Štatistika predstavuje bežne používanú metódu prezentácie informácií, ktorej hlavným cieľom je poskytovať užitočné informácie prostredníctvom čísel. Vo všeobecnosti sa štatistika vzťahuje na numerické dáta. Na základe tohto môžeme štatistiku definovať ako typ informácie získanej prostredníctvom matematických operácií na numerických dátach. Dáta môžu byť zbierané tromi základnými spôsobmi: sčítaním, prieskumami realizovanými na vybraných vzorkách a analýzou administratívnych dát. Každý z daných spôsobov má svoje výhody i nevýhody. Pri výbere správnej metódy zberu dát je potrebné brať do úvahy rôzne faktory. Sčítanie zahŕňa zber dát o každej jednotke/entite v rámci skupiny či populácie. V rámci prieskumu sa zhromažďujú údaje len o vybranej

vzorke subjektov a objektov. Administratívne dáta sú získavané ako výsledok každodennej činnosti (napr. dátum narodenia, adresa a pod.), neskôr sú často podkladom pre prieskum alebo sčítanie. Častým problémom pri používaní štatistických informácií je ich nesprávna interpretácia. Môže byť spôsobená viacerými faktormi, ako napr.: nepochopenie dát, používanie neporovnateľných definícií, zámerná dezinterpretácia informácií. Mali by sme klásť dôraz na dôveryhodnosť zdroja informácií a snažiť sa aj zistiť metódy získavania a spracovávania dát. Faktické informácie sa musia vyznačovať integritou, objektivitou a presnosťou. Pri vyhodnocovaní štatistík je potrebné vziať do úvahy aj možné chyby. Stretávame sa jednak s tzv. výberovými chybami, ktoré vznikajú v dôsledku toho, že pri prieskume sa berie do úvahy len určitá časť/vzorka skupiny a jednak s nevýberovými chybami, ku ktorým dochádza pri prieskume, ako aj pri sčítaní kvôli chybe prístroja, výskumníka či respondenta. Nevýberové chyby môžu byť spôsobené aj neúplnosťou odpovedí alebo stratou získaných informácií. Nevýberové chyby môžeme rozlíšiť podľa povahy na náhodné a systematické.

Knižnice vyhotovujú štatistiky pre rôzne ciele. Používajú ich však skôr na administratívne účely (výročné správy, ročné štatistické výkazy), ako na podporu rozhodovacieho procesu a efektívne plánovanie.

Typy zbieraných a vyhodnocovaných informácií Akademickou knižnicou UK:

- výkaz o akademickej knižnici vo všeobecnosti zahŕňa:
  - informácie o fondoch (knižničné jednotky),
  - informácie o používateľoch,
  - informácie o poskytnutých službách (absenčné a prezenčné výpožičky, rešeršné a informačné služby, študijné priestory, prevádzkové hodiny),
  - doplňujúce informácie (edičná činnosť, vzdelávacia činnosť),
  - informácie o technickom vybavení (počet PC, pripojenie k internetu, licencované zdroje),
  - informácie o pracovníkoch knižnice,
  - finančné informácie,
  - priestory,
- prehľady a štatistiky publikácií a ohlasov,
- využívanie elektronických informačných zdrojov (podľa údajov od dodávateľov).

V ére, keď knižniciam dominovali tlačene dokumenty sa štatistiky takmer výhradne sústreďovali na počty knižničných jednotiek, výpožičiek a registrovaných používateľov. Informácie o nich vychádzajú z nevedeckých (manuálnych) metód zberu a spracovania dát (Čo presne znamená, že kniha nie je na svojom mieste na polici? Je to výpožička? Strata? Nesprávna manipulácia?). Spôľahlivosť a presnosť takýchto informácií je problematická. Údaje sú do veľkej miery nekonzistentné, napr. pri vyčíslení výpožičiek sa neberú do úvahy rôzne výpožičné lehoty a iné lokálne zvyklosti. Knižnice si vedia vypočítať, koľko ich stojí budovanie a prevádzka fondov, ale nevedia presne, ako sú tieto fondy využívané. Výraznú zmenu v tejto oblasti priniesla automatizácia a implementácia knižnično-informačných systémov, automatické sledovanie výpožičiek, čo knihovníkom umožnilo získať relevantnejšie údaje o fondoch a cirkulácii dokumentov. Užitočné informácie poskytujú aj z online katalógy. Údaje o tom, čo a ako používatelia vyhľadávajú, môžu byť veľmi prospešné pre profiláciu fondu, či zameranie podujatí informačnej výchovy.

AK UK	2009	2010
<b>Knižničné jednotky</b>	1 293 614	1 316 348
<b>Registrovaní používatelia (aktívni)</b>	18 751	19 034
<b>- z toho študenti UK</b>	16 313	16 603
<b>Potenciálni používatelia</b>	32 363	32 697
<b>Výpožičky</b>	497 680	1 164 414

Tab. 1 Výber zo štatistického výkazu o akademickej knižnici za roky 2009 a 2010

Dôležitosť konzistencie štatistík a metodiky spracovania môžeme ilustrovať na Tab. 1 obsahujúcej údaje, ktoré Akademická knižnica UK uvádza vo výkazoch školskej a akademickej knižnici pre Ministerstvo školstva, vedy, výskumu a športu SR. Akademická knižnica UK za rok 2010 vykazuje viac ako zdvojnásobený počet výpožičiek oproti roku 2009, pričom počet knižničných jednotiek i čitateľov je relatívne rovnaký. Na prvý pohľad by sme sa mohli potešiť, že študenti začali seriózne študovať (70 výpožičiek na študenta), fond nám cirkuluje (požičaná bola takmer každá kniha). Avšak pri hlbšom skúmaní zistíme, že nárast výpožičiek je spôsobený tým, že sa zmenil spôsob vykazovania výpožičiek a k tradičným absenčným a prezenčným výpožičkám knižničných jednotiek na fyzických nosičoch sa od roku 2010 sem zarátavajú aj prezretia záverečných prác v elektronickej podobe. Záverečné práce sú prístupné online v PDF formáte a každé nahliadnutie (otvorenie súboru) sa započítava ako jedna výpožička. Online záverečné práce predstavujú 825 710 „virtuálnych výpožičiek“, počet reálnych výpožičiek v porovnaní s minulým rokom poklesol, avšak z výkazu to nevidno. V roku 2010 prebehol aj pokus o štatistické vykazovanie licencovaných elektronických zdrojov prístupných v sieťovom prostredí. Zatiaľ nebol vyhodnotený, keďže sa ukázala nejednoznačnosť chápania pojmov a rôznych prístupov, napr. sa zisťoval počet databáz, avšak nebolo definované, ako postupovať pri agregovaných elektronických informačných zdrojoch (EIZ) ako EBSCO či ProQuest – jedna databáza alebo súčet databáz obsiahnutých v EIZ? Ako urobiť sumarizáciu, keď knižnica nerozlišuje medzi EIZ zaobstarávanými konzorciálne a tými, ktoré si knižnica sama predpláca?

EIZ a služby s nimi spojené sa v súčasnosti dostávajú do popredia. Množstvo používateľov ich preferuje a sú aj takí, čo knižnicu fyzicky nenavštevujú a využívajú len online služby. Nástup EIZ dal knihovníkom k dispozícii nové nástroje na zber a spracovanie dát. Vyhodnocovať by sa mali jednak štatistiky poskytované dodávateľmi externých informačných zdrojov, ako aj lokálne generované dáta (logy z proxy serverov, portálov, brán). Oba druhy štatistík ukazujú rôzne formy interakcie používateľov s knižnicou spôsobom, ktorý nebol za „printových“ čias možný. Štatistike vzťahujúcej sa na elektronické informačné zdroje a ich využívanie hovoríme aj e-metrika. Rozoznávame 4 dimenzie e-metriky: externé (dodávateľské) údaje o využívaní EIZ; lokálne dáta; prehľady o používateľoch a fondoch; náhodne alebo pravidelne realizované prieskumy.

V minulosti bolo dosť problematické vyhodnocovanie využívania elektronických informačných zdrojov najmä preto, že dáta od dodávateľov boli navzájom (ale aj medziročne u jedného dodávateľa) často neporovnateľné (sledovali sa rôzne ukazovatele), v nevhodných formátoch, s nízkou informačnou hodnotou. Významný pokrok v tejto oblasti priniesol projekt COUNTER (Counting Online Usage of Networked Electronic Resources), ktorý viedol k vypracovaniu štandardov pre zaznamenávanie a reportovanie využívania EIZ. Riešitelia nadviazali na predošlé aktivity ako ICOLC, ARL e-metrics, NISO Standard Z39.7. V centre pozornosti sú najmä zdroje/kolekcie a ich využívanie (počet stiahnutých plných textov, počet vyhľadávaní a pod.), používateľom ako takým sa pozornosť nevenuje. Výsledkom je určitá analógia k počtu výpožičiek pri tlačených dokumentoch. Takéto

informácie môžu rezultovať do rozhodnutí vedúcich síce k vyššiemu využívaniu EIZ, ale nie nutne aj k lepším a kvalitnejším službám pre používateľov. Benefit je meraný prostredníctvom ceny za použitie (cost per use), nie podľa hodnoty, ktorú to má pre používateľa. Veľa producentov EIZ má nástroje na spracovanie štatistík používajúce COUNTER, napr. 360 Counter, SirsiDynix Directors Station, Scholarly Stats a iné. Ich množstvo a rôznorodosť síce svedčí o záujme o danú problematiku, avšak prináša to problémy v oblasti štandardov a kompatibility. Od roku 2008 sa do praxe zavádza SUSHI (Standardized Usage Statistics Harvesting Initiative). Na objektívnejšie vyhodnotenie využívania EIZ by sa informácie získané na základe štandardov COUNTER, SUSHI a pod. mali dať do súvisu s prieskumami správania a motivácie používateľov. Do úvahy by sa mali brať aj lokálne informácie (návštevnosť a využívanie portálov či webových stránok knižnice). V rámci EIZ sa obvykle stretávame s nasledovnými kvantifikovanými typmi údajov: prístupy (sessions), hľadania (searches), úplné texty (niekedy rozlíšené podľa formátu – html, pdf), abstrakty, stránky (pages), articles, odkázané dokumenty (outward links to documents), dokumenty (documents) a iné.

Akademická knižnica UK pre akademickú obec univerzity zabezpečuje prístup do viac ako 50 000 titulov elektronických kníh a 45 000 online plnotextových časopisov (využívanie pozri Tab. 2). Časť elektronických online časopisov UK získava v rámci participácie na celoslovenskom projekte NISPEZ (Národný informačný systém podpory výskumu a vývoja na Slovensku – prístup k elektronickým informačným zdrojom) garantovaným Centrom vedecko-technických informácií SR. Prístup do týchto zdrojov je z webovej stránky Akademickkej knižnice UK a je – v súlade s podpísanými licenčnými zmluvami – určený výhradne akademickej obci UK. Akademická obec UK má možnosť prístupovať k databázam aj z počítačov mimo UK formou tzv. vzdialeného prístupu (na základe identifikácie cez meno a heslo).

AK UK - 2010	Prihlásenia	Vyhľadávania	Plné texty dokumentov
Portál Web of Knowledge	51 546		
Current Contents Connect	3 274	10 516	
Journal Citation Reports	2 628	3 828	
Web of Science	12 807	13 6804	
Essential Science Indicators	157	1 696	
SCOPUS	22 775	83 048	35 566
ACM	6 415	2 119	2 806
Art Museum Image Gallery	81	253	8 367
Oxford Journals			13 498
EIFL Direct EBSCO	17 266	233 332	23 314
Ebrary	6 432	22 870	7 483
Gale Virtual Refer. Library	2 367	2 628	9 999
IEEE	4 199	2 019	5 021
JSTOR	15 342	71 649	18 933
Knovel	2 523	12 682	2 107
ProQuest	4 973	12 005	2 857
Reaxys	4 857	21 597	
Science Direct	46 426	43 831	149 340
SpringerLink			30 144
Wiley Online	8 625	3 482	21 497

Tab. 2 Využívanie vybraných elektronických informačných zdrojov

Pri hodnotení využívania EIZ je potrebné brať do úvahy rôzne faktory, napr. ako dlho je daný zdroj prístupný, ako sa k nemu používateľ dostáva, vplyv federatívnych vyhľadávacích nástrojov, interakciu používateľov, hodnotu zdroja, rozhranie, atď. Z uvedeného vyplýva, že ani sofistikovanejšie štatistiky (elektronické vs printové) nám nemusia poskytnúť želaný presný obraz použitia zdrojov.

Pri EIZ sa v rámci akademických inštitúcií berie do úvahy aj počet tzv. FTE (Full-time equivalent), celkový počet študentov. S týmto ukazovateľom sa pracuje rôzne. Niektoré akademické knižnice ho zaradili medzi kľúčové výkonnostné indikátory (napr. počet stiahnutí na FTE za rok). V našich podmienkach s týmto ukazovateľom pracujeme skôr pri výpočte ceny EIZ (veľkosť cieľovej skupiny používateľov).

Ku kvantifikovaným údajom využívania EIZ by sme mali pristupovať prezieravo. Veľké čísla pôsobia dobre. Pri interpretácii Tab. 2 by sme si však mohli položiť viacero otázok, napr.: Čo znamená vysoký počet vyhľadávání v rámci EBSCO – zvedavosť používateľov, komplikované rozhranie zdroja, nedostatočnú informačnú gramotnosť? Ako si vysvetliť vyšší počet prístupov ako hľadání (ACM)? Znamenajú nižšie počty prístupov a vyhľadávání (Gale, IEEE), že zdroje sú pre používateľov nezaujímavé? Ale vieme ako početnej cieľovej skupine sú obsahovo určené? Je predsa rozdiel medzi multidisciplinárnym, bibliografickým zdrojom typu WoK, ktorý je určený pre celú akademickú obec a IEEE, ktorý je relevantný pre niekoľko katedier. Ako presne sa vyhodnocuje prístup? EIZ zvyknú mať časový limit na trvanie session, keď používateľ pokračuje v prezeraní dokumentu po prekročení tohto limitu, už sa to ráta samostatne.

Je možné predpokladať, že pri určitých typoch údajov o využívaní dodávateľa EIZ váhajú s ich sprístupnením v obave, že knižnice zrušia predplácanie prístupu k nim. Aj knižnice sa niekedy zdráhajú podeliť o tieto informácie, pretože sa obávajú, že vedenie fakulty či univerzity môže dôjsť k záveru, že prínos knižnice pre inštitúciu klesá. Je veľmi dôležité, aby všetci, ktorým sa štatistika prezentuje, správne chápali význam jednotlivých údajov, vedeli dobre interpretovať dáta i kontext, pre ktorý sú relevantné. Vzhľadom na to, že producenti a agregátori EIZ vykonávajú väčšiu časť práce spojenej so zberom, organizáciou a uchovávaním digitálnych informácií, knižnice by mali v rámci svojej činnosti zamerať svoju pozornosť, aktivitu a snahu na vzdelávanie a výskum. Od knihovníkov sa očakávajú výborné zručnosti v oblasti vyhľadávania informácií, informačného prieskumu. Mali by sprostredkovať informácie používateľom, pomáhať im vyhodnotiť relevanciu, dôveryhodnosť, užitočnosť, atď. nájdených informácií. Určitým spôsobom by mali na seba prevziať aj časť zodpovednosti za výsledky vedecko-pedagogického procesu. EIZ a online služby vyžadujú, aby knihovníci viac spolupracovali so širokou škálou odborníkov ako sú napr. archivári, grafickí dizajnéri, programátori, psychológovia a pod. Na kompetencie a výkony knihovníkov či informačných profesionálov sa v súčasnosti kladú naozaj veľké nároky. Tradičné merania výkonov, štatistické ukazovatele však nezachytávajú tieto nové úlohy a zodpovednosti.

E-metrika hrá dôležitú úlohu pri riadení a správe EIZ a finančných rozhodnutiach. Avšak mala by byť správne integrovaná do celkového hodnotenia knižnice, ktoré odráža, ako knižnica splňa inštitucionálne ciele.

Financovanie knižníc nebolo nikdy optimálne, súčasná ekonomická kríza situáciu ešte zhoršila. Nanešťastie práca knihovníkov je často neviditeľná pre tých, ktorí rozhodujú

o financiách. Knižnice zlyhávajú pri prezentovaní svojej hodnoty, nevedia sa „predať“. Vnímanie knižníc verejnosťou (resp. akademickou obcou v prípade akademických knižníc) sa zväčša pohybuje v rovine poskytovateľov informácií. Je paradoxom, že nástup internetu, ako aj konzorcionálne projekty financované z iných ako inštitucionálnych zdrojov i možnosť využívať EIZ prostredníctvom tzv. vzdialeného prístupu danú situáciu do značnej miery zhoršujú, pretože mnohí používatelia majú pocit, že informácie sú zadarmo, keďže k nim majú relatívne bezprostredný online prístup. Takýto postoj robí z informácií „neznačkovú“ komoditu, ktorou bez pridanej hodnoty ťažko oslovíme používateľov. Môžeme si za to do istej miery aj my sami. Nezdôrazňujeme používateľom, že na všetky elektronické informačné zdroje, ku ktorým majú „zadarmo“ prístup v sieťovom prostredí, sa v skutočnosti vynakladajú nemalé finančné prostriedky. Tieto elektronické informačné zdroje sú často také drahé, že nútia k spolupráci, k spájaniu finančných prostriedkov viacerých inštitúcií, či k hľadaniu iných ako rozpočtových zdrojov. Ak nezdôrazníme, že databázy sú pre našich používateľov „bezplatné“ len preto, že ich za nich predpláca inštitúcia či knižnica, môžeme sa stať obeťami názoru, že knižničná činnosť je zbytočná. Veď všetko potrebné je „voľne dostupné“ na webe. Nemôžeme očakávať automatickú podporu, musíme si ju aj zaslúžiť – a to nielen dobrou každodennou prácou, ale aj príslušnou propagáciou. A práve tu nám môže poslúžiť štatistika. Tú možno využiť nielen smerom navonok – ako prezentáciu výkonov, sledovanie využívania, akvizičné rozhodnutia, ale aj smerom dovnútra – na optimalizáciu knižničných procesov.

### **Bibliografické odkazy**

ANDERSON, Elise. 2005. Maximizing the value of usage data. *Against the Grain*. 2005, 17 (5), s. 20-22.

DUGAN, R.E., HERNON. 2002. Outcomes assessment: Not synonymous with inputs and outputs. *The Journal of Academic Librarianship*. 2002, 28 (6), s. 376-380. doi:10.1016/S0099-1333(02)00339-7

HILLER, S., SELF, J. 2004. From measurement to management: using statistics wisely in planning and decision making. *Library trends*. 2004, Vol. 54 No. 1, s. 129-155.

McDONALD, John. 2007. Usage statistics and information behaviors: Understanding user behavior with quantitative indicators [online]. Paper presented at Understanding the Data Around Us: Gathering and Analyzing Usage Data, NISO Usage Data Forum, November 1-2, in Dallas. [cit. 15-04-2011]. Dostupné na internete: <<http://niso.kavi.com/news/events/niso/past/usage07/usage07mcdonald.pdf>>.

PFEFFER, Jeffrey, SUTTON, Robert I. 2006. *Hard Facts, Dangerous Half-Truths And Total Nonsense: Profiting From Evidence-Based Management*. Boston : Harvard Business Press, 2006.

WAWRZASZEK, Susan V., WEDAMAN, David G. 2008. The academic library in a 2.0 world [online]. *Research Bulletin*. 2008, no. 19. [cit. 15-04-2011]. Dostupné na internete: <<http://www.educause.edu/ECAR/TheAcademicLibraryina20World/163206>>.

WHITE, Larry Nash, BLAKENSHIP, Emily F. 2007. Aligning the assessment process in academic libraries for improved demonstration and reporting of organizational performance. *College & Undergraduate Libraries*. 2007, 14 (3), s. 107-119.

WHITEHEAD, Derek. 2007. Why have statistics? ... And what? [online]. Paper presented at the CEIRC Forum on Usage Statistics for E-Resources, in Sydney, Australia. [cit. 15-04-2011]. Dostupné na internete:  
< <http://www.caul.edu.au/content/upload/files/datasets/usagestatistics2007whitehead.pdf> >

*Výkaz o školskej a akademickej knižnici za rok 2009 : Akademická knižnica UK.* Bratislava : [s. n.], 2010. 2 s. Interný materiál.

*Výkaz o školskej a akademickej knižnici za rok 2010 : Akademická knižnica UK.* Bratislava : [s. n.], 2011. 2 s. Interný materiál.

*Výročná správa o činnosti Univerzity Komenského v Bratislave za rok 2010.* Bratislava : Univerzita Komenského, 2011.