



**ÚSTAV INFORMAČNÍCH STUDIÍ A KNIHOVNICTVÍ  
FF UK V PRAZE**

**Lucie Vavříková**

# **Úvod do scientometrie**

**Verze 1.0**

**Praha**

**Říjen 2008**

## 1 ÚVOD DO SCIENTOMETRIE

„In science, the quantity and quality are correlated.“ [Dennis 1954]

Jak napovídá úvodní citát, ve vědě lze měřit kvalitu a kvantitu [GODIN, 2006]. *Scientometrie* je věda, která se tímto měřením zabývá, snaží se nalézat metody a metriky pro určení jednoho či druhého aspektu, obecněji se také vžil pojem měření výkonnosti vědy. Ve své samotné podstatě sleduje a hodnotí komunikaci ve vědě, neboť všechna měření jsou odvozena od interakcí mezi jednotlivými elementy scientometrie, touto základní interakcí je citace. Její název je odvozen z *metrein* – měřit (řec., lat.), *scientia* – znalost, vědění. *Scientometrie* přímo navazuje a vychází z informační vědy a používá mj. její metody, pro správné pochopení je ji třeba vsadit do kontextu jednotlivých metod *bibliometrie* a *informetrie*. Prakticky je *scientometrie* dnes nejvíce využívána pro hodnocení vědy pro rozdělování finančních zdrojů ve vědě.

### 1.1 KOMUNIKACE VE VĚDĚ

Pro rozvoj a vývoj vědy a výzkumu je nezbytná vzájemná komunikace vědeckých poznatků – komunikovány jsou *vědecké informace*, a to v různých formách vědeckých výstupů; nejběžnějším a nejpoužívanějším z nich je článek v časopise, k dalším obvyklým patří konferenční příspěvek, monografie, review až po postery a jiné typy dokumentů. Podstatou komunikace je tedy publikace, která po svém zveřejnění má jistý dopad a vliv ve světě vědy. Tento dopad a vliv se odráží v citacích a jejich míře. Odborné literatura je pak uvažována jako rozsáhlá síť, kde jsou jednotlivé prvky spojeny pomocí citací a dalších vztahů (spoluautorství).

Pro úplnost se ještě vraťme k vědeckým informacím a k jejich vlastnostem, které se významně promítají v komunikaci a projevují se v metodách scientometrie. Specifické vlastnosti vědeckých informací jsou velmi podstatné pro formulování teorií, metod a jejich způsobům aplikace. Scientometrie např. podle vlastností vědeckých informací dokáže charakterizovat jednotlivé vědecké disciplíny – typický je pro jednotlivé disciplíny koeficient stárnutí informace.

*Vlastnosti vědeckých informací* [KÖNIGOVÁ, 2002]:

- přesnost, objektivnost, pravdivost, platnost, jednoznačnost, účelnost, společenská hodnota atd. (tím se vědecké informace odlišují od ostatních druhů sociální komunikace)
- stárnutí - snižování hodnoty v závislosti na čase
- sémantická povaha - zachycené smysluplným souborem znaků
- nezávislost na jazyku a nosiči - tentýž obsah může mít různé formy
- neaditivnost, nekomutativnost, neasociativnost
- rozptyl informací v časopisech
- kumulativnost nebo možnost redukce
- nezávislost na tvůrcích

Jak jsme uvedli, publikace je jen jednosměrnou formou komunikace, to, co umožňuje získávat celý obraz, jsou tzv. reference, v češtině nejčastěji uváděné za dílem jako seznam použité literatury a jeho retrospektivní propojení s uvedenými díly, které tak mohou být opatřena výčtem citací. Díky informačním a komunikačním technologiím – relačním databázím, automatickým mechanismům sběru, uložení a vyhledávání, a zejména sofistikovaným technologiím propojování záznamů mezi sebou dnes máme k dispozici silné nástroje a získáváme tak mnohem jednodušeji než dříve schéma vědecké komunikace, ze které lze dále usuzovat a provádět mnohé analýzy, které ať už jsou náplní scientometrie nebo přesahují právě do jiných oborů, jako je sociologie, psychologie či historie jednotlivých disciplín.

## 1.2 HISTORIE

Scientometrie se zejména v posledních letech dostala na výsluní pozornosti napříč společnostmi, ať už se jedná o sféru akademickou, vědeckou, průmyslovou nebo na nejvyšší úrovni sféru vládní. Důvodem je zejména vliv výsledků scientometrie pro rozdělování finančních zdrojů ve vědě, dále exponencionální nárůst vědecké literatury, kdy je velmi těžké v nepřeberném množství zdrojů nalézt kvalitní vědecké výstupy, v neposlední řadě pak rozvoj scientometrie souvisí s rozvinutými možnostmi informačních a komunikačních technologií.

Za zakladatele jsou dnes v mezinárodním kontextu považováni Derek J. de Solla Price a zakladatel citačních rejstříků Eugene Garfield [Scientometrics, 2007]. Historicky můžeme první pokusy nalézt na konci 19. století v oblasti práva (mnohem dříve než pro oblast vědecké literatury! - právo je považováno za aplikovaný obor), začaly být praktikovány bibliometrické metody [SHAPIRO, 1999]. V kontextu bibliometrie vzpomeňme francouzské a belgické dokumentační snahy, ve Francii se objevují první náznaky kolem r. 1880, v Belgii je významným datem r. 1895 a založení Mezinárodního bibliografického ústavu Paulem Otletem a Henrim Lafontainem.

V r. 1903 se byla provedena kvalitativní a kvantitativní analýza v oblasti psychologie, byl vytvořen soupis přibližně 200 psychologů s dokumentací jejich afiliací a konkrétních oborů bádání [GODIN, 2006], v této analýze je vyhodnocen nejlepší vědec, je provedeno porovnání států. V r. 1926 je definován Lotkův zákon o vztahu počtu autorů k počtu děl, která vytvářejí, 1935 pak George Kingsley Zipf v rámci kvantitativní lingvistiky popsal pravidla frekvence slov v textu (Zipfův zákon) a v r. 1948 Bradfordův zákon navázal na Lotku a definoval tzv. jádro časopisů, tj. to, že většina článků vychází jen v malém počtu časopisů.

Další významný moment a zároveň posun k samostatné vědě datujeme k r. 1944, kdy Harvey C. Lehman popsal vztah mezi kvalitou a kvantitou v oblasti vědeckého publikování, 1952 ho následoval Dennis, který rozvinul jeho práci a analyzoval vliv věku vědce na tyto dvě charakteristiky, zároveň vyvrátil Lehmanovy teze o platnosti přímé úměry mezi nimi, ale zároveň potvrdil jejich vzájemný vztah, když konstatoval: „In science, the quantity and quality are correlated“. Na počátku druhé poloviny 20. století tak pronikají do současné odborné literatury popisy a definice zmíněných metod zkoumání a etablojí se jako metody bibliometrie.

V r. 1955 přichází Eugene Garfield se svou teorií impakt faktoru<sup>\*</sup>, který byl původně definován jako indikátor vypočtený z citačních rejstříků, které sám Garfield vyvíjel, pro výběr časopisů v oblasti genetiky (*Genetics Citation Index*). Garfield v podstatě formuloval myšlenky hodnocení časopisů podle toho, jaký mají vliv a dopad ve vědě podle počtu získaných citací. Pro účel vytváření rejstříků, které bylo velmi náročné vzhledem k velkému objemu manuální práce, založil s pověstnými pár dolary v kapse v r. 1958 *Institute for Scientific Information* (dnes existuje v rámci mediálního giganta a producenta Thomson Reuters). V r. 1961 byl poprvé definován *Science Citation Index* (dále SCI), citační rejstřík pro soubor exaktních věd, první SCI byl pak vydán 1963 [SCHWARZ, 2003]. O deset let později vznikl rejstřík sociálních věd *Social Sciences Citation Index* (dále SSCI) a v r. 1979 *Arts and Humanities Citation Index* (dále AHCI).

Významnému rozvoji scientometrie přispěli také vědci z východního bloku, zejména když v r. 1959 ruský vědec Nalimov se svými kolegy v článku *Scientific and technical information as one of the tasks of cybernetics*<sup>†</sup>, komentovali nárůst vědy, nárůst potřeby času při hledání odborné literatury, nutnost expertních/matematických metod [GRANOVSKY, 2001]. Nicméně režim v této části světa samozřejmě zasahoval do „názoru“ na vědu a do vědy o vědě, která taktéž byla státně řízena a hlídána. I přes nastalé problémy po vydání článku se však rozvoj nezastavil, došlo k vydání dalších publikací, mj. s velkým úspěchem v r. 1969 článek, ve kterém byl poprvé definován pojem scientometrie, resp. naukometrie.

V šedesátých letech zaznamenáváme obrovský nárůst literatury publikované k tématu scientometrie, objevuje se velmi významné jméno Derek de Solla Price, který pracoval na studii exponencionálního nárůstu vědy a citační aktivity odborné literatury. Publikoval několik prací popisujících klíčové body scientometrických analýz, včetně vzorů komunikace mezi vědci a historií samou o sobě a vědy samotné. Solla Price je považován za zakladatele scientometrie, od r. 1984 se uděluje cena D.S. Price, v r. 1989 ji dokonce získal český fyzik Jan Vlachý, který velkou část svého díla věnoval právě

---

\* ang. orig. Impact Factor, v textu je používáno počeštěné spojení impakt faktor.

† VLEDUTS G.E., NALIMOV V.V., STYAZHKIN N. I. *Scientific and technical information as one of the tasks of cybernetics. Uspěchy fyzických nauk*. 1959, T. 41, no. 69, s. 13-56.

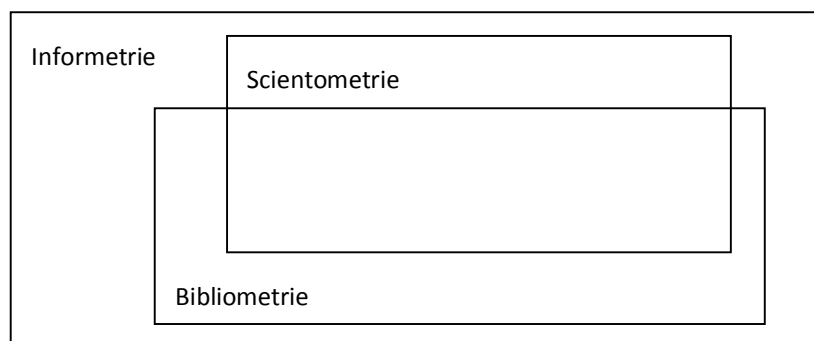
citačním analýzám. Cenu v současné době uděluje profesní sdružení *International Society for Scientometrics and Informetrics*<sup>\*</sup>, která vznikla v r. 1993 na mezinárodní konferenci bibliometrie, informetrie a scientometrie<sup>†</sup>, která se dodnes uskutečňuje každé dva roky. Na sklonku sedmdesátých let začíná vycházet časopis *Scientometrics*<sup>‡</sup>, který se zabývá kvantitativními aspekty vědy o vědě, komunikací ve vědě a politikou vědy a výzkumu. Vydává ji Maďarská akademie věd, zakladatelem a dodnes redaktorem je významný vědec Tibor Braun.

Scientometrii velmi ovlivnil nástup informačních a komunikačních technologií, po nástupu databází lze sledovat změnu paradigmatu scientometrických analýz, vznikají nové technologie bibliometrického výzkumu včetně metody mapování citací. Na počátku 90. let došlo k nejrazantnějšímu rozvoji těchto technologií, hned v r. 1990 byla spuštěna webová verze citačních rejstříků. Web přinesl scientometrii předtím lidsky nemožné věci: počty stažených článků, Google Page Rank (číselná hodnota určující důležitost stránky na webu), analýzy používání webu. Ze scientometrie se tak také začíná vydělovat další dílčí disciplína, webometrie, která omezuje svá zkoumání na webové prostředí, v r. 1998 je dokonce navržen *web impact factor*.

Z nejnovějších významných událostí ve světě scientometrie jmenujme spuštění citační a abstraktové databáze Scopus<sup>§</sup> v r. 2004 nizozemským nakladatelstvím Elsevier a definování nového indikátoru H-Index v r. 2005, který se velmi rychle rozšířil a získal velkou oblibu.

## 2 INFORMETRIE, BIBLIOMETRIE A SCIENTOMETRIE

Scientometrie, informetrie a bibliometrie jsou si velmi blízké, bývají možná i právem někdy zaměňovány, to však může být způsobeno právě tím, že se navzájem prolínají, spolupracují spolu, používají stejné metody, jednu si dnes bez druhé nelze představit. V pramenech se také objevuje několik různých pojetí jejich vzájemného vztahu, vymezení každé z nich není zcela jednoznačné. Jejich vztah je naznačen na obr. č. 2:



Obr. 1 – Vzájemný vztah informetrie, scientometrie a bibliometrie

Zastřešující je informetrie, jak naznačuje její název, zabývá se informacemi, respektive informačním procesem a měřením informačních toků. Bibliometrie a scientometrie se navzájem velmi překrývají, mají totožné vědecké metody, dělají obdobné analýzy, nicméně bibliometrie je omezena na zaznamenané informace a jejich bibliografie (řekněme publikace obecně), zatímco scientometrie je vymezena sociálním aspektem vědy, zkoumá pouze publikace, resp. vědecké výstupy ve vědecké komunitě. Bibliometrie je „starší“, v podstatě je scientometrie jen její nadstavbou a doplňkem; scientometrie obecně ke svému zkoumání využívá zejména bibliometrických metod. Pojmy bibliometrie a

\* dostupná na <http://www.issi-society.info/>

† International Conference on Bibliometrics, Informetrics and Scientometrics

‡ *Scientometrics : an international journal for all quantitative aspects of the science of science, communication in science and science policy*. Akademiai Kiado Rt. 1978- , vol. 1, no. 1- . [Budapešť] : Maďarsko, 1978- . ISSN 0138-9130.

§ dostupné na <http://www.scopus.com>

scientometrije uvedl Pritchard a Nalimov s Mulčenkem takřka ve stejný čas (1969), informetrie si na svou definici počkala dalších 10 let.

### **Scientometrie**

**Scientometrie**, v angličtině „Scientometrics“, se obecně řečeno zabývá měřením vstupů a výstupů ve vědě a měřicími procedurami ve vědě [KÖNIGOVÁ, 2001]. Z pohledu sociologického se jedná o studium kvantitativních aspektů vědecké komunikace (publikace, citace). Jejím cílem je vyvinout indikátory intelektuální a sociální organizace vědeckých disciplín za pomoci vzájemných vztahů mezi autory a texty [LEYDESDORFF, 2001]. Scientometrie provádí matematické a statistické analýzy vědeckého výzkumu (častěji v tzv. tvrdých vědách) [KTD, 2003]. Pomocí scientometrie lze určit vzory a identifikovat charakteristiky jak autorů, tak dokumentů. Metodami scientometrie lze propojit informace o institucích na úrovni výzkumných skupin s vývojem na úrovni jednotlivých disciplín a oborů. Jeden z nejdůležitějších cílů scientometrie je zkoumat kvantitativní vztahy produkce, diseminace a impaktu/vlivu vědeckých informací [VINKLER, 2008].

Pokud uvážíme dnešní ekonomické aspekty hodnocení vědy, lze definici scientometrie ještě o kousek posunout, scientometrie zkoumá vědeckou produktivitu a její výslednou prospěšnost [SOUČEK, 2006], současný výzkum využívá jak bibliometrické tak ekonomické analýzy [DIODATO, 1994], věda je uvažována jako ekonomická aktivita [TAGUE-SUTCLIFFE, 1992].

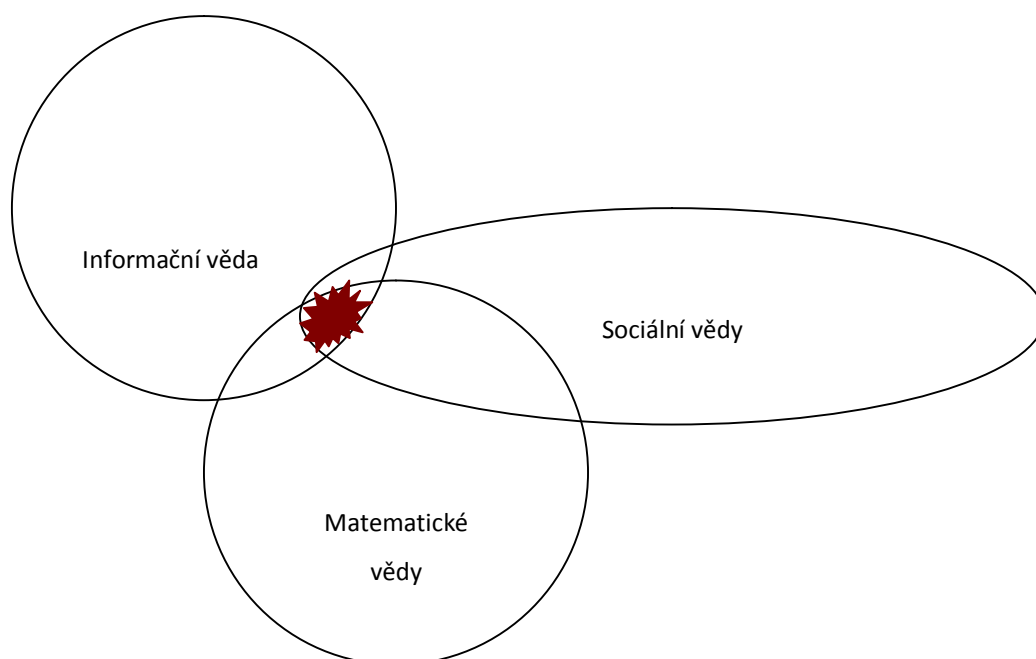
Scientometrii můžeme rozdělit na několik dílčích disciplín, zejména podle toho, na jaké výstupy se orientuje, jaký je jejich cíl. *Strukturální scientometrie* se orientuje na zkoumání mapování epistemologické struktury vědy, *dynamická scientometrie* vytváří sofistikované modely vývoje vědy, jejího nárůstu, zastarávání, citačních procesů aj. *Evaluativní scientometrie* se zaměřuje na vývoj indikátorů, které by charakterizovaly výkonnost vědy a výzkumu na různé úrovni agregace [GLÄNZEL, DEBACKERE, 2004]. Tato specifická oblast zkoumá *scientometrické aspekty scientometrických entit* tak, aby mohla kvantifikovat relativní výkon určitých *organizací*. Hlavní pole studia jsou komparativní studie informační produkce a vlivu informací hodnocených organizací. Jako *scientometrickou entitu* chápeme tematicky, institucionálně nebo organizačně vymezenou entitu, kterou můžeme charakterizovat alespoň jedním *scientometrickým prvkem*. *Scientometrické aspekty* jsou kvantifikovatelné charakteristiky jevů ve vědě nebo vědeckém výzkumu, které však nejsou součástí primárního zájmu specifických vědeckých disciplín. Nejvíce relevantní ve vědě v tomto ohledu jsou aspekty informačních procesů. Kvantifikace obvykle (ale ne vždy) značí aplikaci statistických metod [VINKLER 2001]. Nejvyužívanější je evaluativní scientometrie na mikroúrovni, tedy u hodnocení jednotlivých autorů.

D. J. de Solla Price, zakladatel scientometrie, se výrazně zabýval oblastí historie vědy a vědy o vědě, z kvantitativních analýz odvozoval tvrzení o exponenciálním charakteru vývoje či rozložení zkoumaných jevů. Mezi nejznámější patří zákon o exponenciálním růstu počtu dokumentů a tvrzení, že polovina vědeckých článků je publikována druhou odmocninou počtu všech autorů [KTD, 2003].

Scientometrie ke svému zkoumání využívá zejména bibliometrických metod (publikační analýzy), vlastních metod (citační analýzy) a jiných. Scientometrie je také disciplínou, která úzce souvisí s dalšími obory, vzhledem k její povaze „měření“ je nezbytné využití matematické statistiky, pro práci s odborným textem a potažmo jazykem je velmi blízká s kvantitativní lingvistikou a s textovou analýzou, interpretace výsledků zasahuje do sociologie, psychologie, teorie komunikace a dalších sociálních oborů. Na obrázku č. 2 je zachyceno přibližné schéma, které vyjadřuje zařazení scientometrie v rámci ostatních věd.

---

\* Nacke, O. (1979). Informetrie: eine neuer Name für eine neue Disziplin. Nachrichten für Documentation, 30(6), 219-226.



**Obr. 2 – Průnik věd, na kterém leží scientometrie**

Scientometrii a to, jak jsou jednotlivé indikátory citačních analýz konstruovány, použity, interpretovány nebo teoreticky založeny, můžeme nahlížet v kontextu pěti základních disciplinárních přístupů [MOED, c2005].

*Fyzický přístup* přináší do scientometrie zákony (bibliometrické, citační), definuje indikátory.

*Sociologický přístup* uvažuje vědecké publikování a citování jako druh, nahlíží na vše jako na sociální akt a podrobuje je zkoumání. Na úrovni mikrosociologické je možné např. zkoumat život vědce či jeho vědeckou kariéru.

*Psychologický přístup* se snaží odhalit psychologii citování, zkoumá např. motivy citování.

*Historický přístup* přináší kognitivní dimenzi odborných aktivit (popisuje vývoj myšlenek aj.) nebo se snaží zjišťovat sociální, ekonomické a institucionální podmínky, které podmiňují vědeckou aktivitu.

*Informačně-vědní přístup* vnáší pohled ze strany teorie informace a jejích konceptů, půjčuje si elementy fyzického a sociologického přístupu.

V dnešní době využívá scientometrie zejména data dvou komerčních produktů, databáze citačních rejstříků dnes sjednocená do platformy ISI Web of Science (Thomson Reuters) s více jak čtyřicetiletou historií a konkurenční citační databáze Scopus od nizozemského vydavatele Elsevier spuštěná před 5 lety. S rozvojem technologií je však stále snazší pro ostatní producenty zabudovávat do svých databází a dalších informačních a znalostních produktů prvky scientometrických analýz, zejména sledovat vzájemné citace. Velmi diskutován je např. nástroj Google Scholar. Z dalších produktů zahrnující tyto nástroje je databáze PubMed, CSA Illumina aj.

### **Informetrie**

Informetrie je ze zmiňovaných třech dílčích disciplín nejširší, její povaha je nejvíce teoretická a zajišťuje takto metodologický základ zbývajícím dvěma. Informetrie se zabývá měřením toku informací, hodnocením informačního procesu. Jedná se o vědní obor, který používá matematicko-statistických metod k popisu a analýze informačních jevů za účelem hledání jejich zákonitostí [KTD, 2003]. Jako součást informační vědy se zabývá především kvantifikací informace, kvantitativní analýzou informačních toků a dokumentů apod.; její výsledky slouží např. pro analýzu kvantitativního růstu literatury, účinnosti informačních systémů či role informací ve vědecké komunikaci. Informetrie je metoda zkoumající a kvantitativně měřící vznik, oběh a působení informací v jakékoli oblasti společenského života jednotlivce nebo společnosti. Usiluje o vyjádření hodnoty informačního sdělení. [KÖNIGOVÁ, 2001].

## **Bibliometrie**

Výraz bibliometrie bývá velmi často používán jako souhrnný pojem, autoři často informetrii nebo scientometrii od ní nerozlišují. Pravděpodobně je to tím, že převážná část metodologického aparátu je definována jako bibliometrická.

Bibliometrie se zabývá studiem kvantitativních aspektů produkce, rozšiřování a užití zaznamenaných informací. Vytváří si matematický model a teorie procesů a poté je aplikuje pro odhadování vývoje a pro rozhodování [TAGUE-SUTCLIFFE, 1992]. Bibliometrie pomocí kvantitativních analýz, statistik a datových vizualizací zjišťuje charakteristiky referencí, citací, autorů, časopisů, institucí, slov, klíčových slov, klasifikačních kódů atd. Typickou a jednou z nejstarších je bibliometrická analýza kvantitativního růstu literatury. Do bibliometrie dnes bývá velmi často také zařazována oblast výzkumu informačních potřeb (uspokojení nebo neuspokojení požadavků uživatelů), neboť bibliometrie tvoří jeho valnou část, jako např. hodnocení intenzity využívání knihovních a informačních fondů.

Jak bylo již výše zmíněno, máme tři bibliometrické zákony Lotkův, Bradfordův a Zipfův, mezi nepoužívanější bibliometrické metody řadíme publikační a citační analýzy.

*Lotkův zákon*, známý jako zákon měření vědecké produktivity [KÖNIGOVÁ, 2001], který byl definován při srovnávání publikační činnosti autorů v oblasti chemie a fyziky [SOUČEK, 2006], říká, že se stoupajícím počtem publikovaných článků klesá počet autorů píšící největší poměr množství článků [Scientometrics, 2007]. Pokud  $a_1$  autorů publikuje jeden článek, pak  $a_n$  bude počet autorů, kteří publikují  $n$  článků. [SOUČEK, 2006]:

$$a_n = a_1 / n^2$$

*Zipfův zákon* neboli frekvenční analýza respektive „rozptyl slov v textu“ [KÖNIGOVÁ, 2001] byl definován Georgem Kingsley Zipfem, americkým psychologem a lingvistou. Zipfův zákon vyjadřuje skutečnost, že základ lexiky vytváří relativně malý počet silně frekventovaných slov. Zipf slova uspořádal podle frekvence ( $f$ ) výskytu, slovo s nejvyšší hodnotou výskytu získalo  $r=1$ , následující  $r=2$  atd. Jsou-li slova dlouhého textu seřazena podle pořadí klesající četnosti jejich výskytu, potom součin pořadí ( $r$ ) a četnosti ( $f$ ) pro každé slovo textu bude přibližně táž konstanta ( $C$ ), která závisí na délce textu.

$$r * f = c$$

Na základě Zipfova zákona se konstruovaly také slovníky nevýznamových (stop) slov, které jsou základem pro automatické vyhledávání.

*Bradfordův zákon*, je zákon rozptylu informací, nebo rozptylu relevantních článků v časopisech [KÖNIGOVÁ, 2001]. Definoval jej Samuel Clemens Bradford, americký knihovník, jeden ze zakladatelů bibliometrie. Tento zákon se prakticky může použít např. pro tvorbu a vyváženost knihovních fondů. Obdobně jako Lotkův zákon definuje nepřímou úměru mezi časopisy a relevantními články v nich - maximální počet relevantních článků je soustředěn v minimálním počtu časopisů tvořících tzv. jádro. Jádro oboru pak tvoří nejužší skupina časopisů, ve kterých vyšlo nejvíce relevantních článků.

## **Webometrie**

S nástupem počítačů, nejen že se díky obrovskému výpočetnímu, propojovacímu a úložnému potenciálu scientometrie velmi rozvíjí, dochází také k transformování bibliometrických metod do tohoto prostředí a formuje se *kybermetrie*, nebo také *webometrie*. Lze je chápat jako totožné, pokud bychom je chtěli od sebe alespoň nějak odlišit, kybermetrie je určena obecně aplikací do kyberprostoru, navíc první časopis zabývající se touto oblastí nese název *Cybermetrics*, webometrie se pak spíše aplikuje v rámci World Wide Web, jinak zjednodušeně označované jako internet nebo web. O obou se začíná hovořit v roce 1997. Vzhledem k dnešnímu vymezení většiny služeb právě v rámci internetu, respektive využívání zdrojů skrze webové rozhraní, na kterém jsou analýzy založeny, bude zde souhrnně použit termín webometrie.

*Webometrie* se zabývá kvantitativními aspekty vytváření a užití informačních zdrojů, struktur a technologií celého internetu, přičemž využívá bibliometrického a informetrického přístupu. Web může být v jistých aspektech nahlížen stejně jako klasická „papírová“ komunikace, kde tradiční informační entity a citace jsou nahrazeny webovými stránkami a odkazy. Výrazným rozdílem proti klasickému prostředí je možnost oboustranného propojení a samozřejmě to, že ve webovém prostředí neexistuje kontrola kvality [BJÖRNEBORN, INGWERSEN, 2004]. Webometrie je stále ještě v experimentálním stádiu, je zkoumáno, zda bibliometrické metody mohou v tomto prostředí být aplikovány [NORUZI, 2006].



Zkoumat lze ve všech složkách digitálního prostředí, tj. v sítích jednotlivých domén, intranetu; webometrie provádí hodnocení sítí podle objemu informačních kolekcí, počtu www stránek, podle četností vstupu, času odezvy. Provádí klasifikaci stránek podle způsobu prezentace, vytváří statistiky užívání, hodnotí produktivitu autorů [SOUČEK, 2006].

### 3 CITAČNÍ ANALÝZY

Tato kapitola se pokusí přiblížit základní metodu zkoumání ve scientometrii: citační analýzy, jejich účel, využití a omezení. Ačkoliv je kapitola pojmenována ve stručnosti „Citační analýzy“, seznámíme se nejprve s analýzami publikačními, které jsou vlastním základem analýz citačních. Citační a publikační analýzy jsou základní bibliometrické metody, které umožňují měření a tím vzájemné srovnávání v oblasti vědy a výzkumu.

Uvést publikační analýzy je nezbytné proto, že pokud chceme získat kompletní obraz vědy např. v daném státě, je nezbytné se zaměřit již na základní úroveň, tj. přímo na samotné publikování, ještě než budeme zkoumat další komunikaci díla v rámci vědy. Pro důslednou interpretaci výsledků citačních analýz je nutné přihlídnout k výsledkům publikační analýzy. Např. pokud bychom z citační analýzy usoudili, že v určitém oboru nejsou v daném regionu kvalitní vědci, neboť nejsou citováni, je důležité nejprve zjistit, zda-li se v daném regionu v daném oboru vůbec publikuje. Citační a publikační analýzy nejsou vždy vymezeny zvlášť, často jsou v rámci analýzy citační provedeny analýzy publikační.

*Publikační analýza* je matematicko-statistická bibliometrická metoda, která se zabývá kvantitativním měřením produkce publikací. Publikační analýzou nejčastěji vyhodnocujeme [INGWERSEN, LARSEN, 2007]:

- geografickou oblast
- vědní oblast
- časovou periodu
- typ vědecké literatury
- autora v oboru nebo zemi
- časopisy v oboru
- instituce

Publikační analýzou stanovujeme např. jaké typy publikací jsou typické pro danou oblast [INGWERSEN, LARSEN, 2007]:

- přírodní vědy a medicína
  - o recenzované články
  - o článek typu review
  - o dopisy a poznámky\* (typické zejména pro fyziku a obdobné oblasti)
- technické a počítačové vědy
  - o recenzované články
  - o příspěvky ve sborníku (recenzované)
  - o patenty
- společenské a humanitní vědy
  - o knihy (antologie a monografie; recenzované a vydané)
  - o recenzované články

*Citační analýza* je matematicko-statistická bibliometrická metoda, která kvantifikuje vztahy mezi autory, dokumenty a vědními obory na základě bibliografických citací a bibliografických referencí. Konkrétně citační analýzy zkoumají

---

\* v angl. orig. letters and notes



citovanost dokumentů, četnosti citací v dalších pracích apod. [KTD, 2003]. Analýzy se zabývají výzkumem citačních vazeb a citací. Jedná se o zkoumání témat, směrů, množství a vazeb, jsou vytvářeny např. citační sítě a grafy [JANSSENS, 2006]. Výsledkem citační analýzy může také být určení hodnoty scientometrického indikátoru [SMELSER, BALTES, 2004]. Analýzou citačních rejstříků je možné určit a popsat směr budoucího vývoje vědních disciplín a mapovat vědu [JANSSENS, 2006]. Pomocí citačních analýz se měří tzv. viditelnost v oblasti vědy\*.

*Citace* je znak vědecké komunikace, je znakem užití informace. Citace navzájem propojuje dokumenty, pomocí nich lze vymodelovat odbornou literaturu jako velkou síť dokumentů, které jsou spojeny citacemi. Citace jsou dvojího druhu: citace a reference. Formálně reference ve vědecké komunikaci označuje použití informace, citace vyjadřuje příjem tohoto použití ve vědecké komunikaci.

Citační analýzy měří a indikují zejména:

- počet citací dokumentu
- vztahy dokumentů
- vzájemnou obsahovou souvislost
- nepřímou obsahovou souvislost (více dokumentů citovaných jedním dokumentem, jeden dokument citovaný více dokumenty)
- jádro oboru
- kvantifikované profily oboru a prognostické odhady

Výsledky citačních analýz lze interpretovat v několika směrech, lze z nich usuzovat na [MOED, c2005]:

- výkonnost vědy
- odbornou kvalitu
- vliv
- dopad

Specifickým druhem citačních analýz jsou takzvané *kocitační analýzy*. Určují podobnost mezi dvěma elementy. Pokud jsou elementy A i B oba citovány elementem C, lze usuzovat na jejich vzájemný vztah, i když na sebe přímo neodkazují. Pokud jsou elementy A a B oba citovány více dalšími elementy, mají silnější vzájemný vztah; čím více jsou elementy společně citovány, tím silnější jejich vztah je. Kocitace byly poprvé navrženy jako základní metrika k charakterizování podobnosti mezi dokumenty. V databázích si širší aplikace této metriky lze povšimnout jako vyhledávání podobných dokumentů (related documents search), někdy nazýván obecně jako „*pattern search*“, vzorové vyhledávání. Podobnost se stanovuje nejen na základě autorů, ale také např. klíčových slov.

Jako aplikovaná metoda má citační analýza význam pro optimalizaci informačních toků a pro profilování knihovních fondů; je také základem citačního mapování vědy pomocí konstrukce citační sítě [KTD, 2003]. Citační analýzy jsou dnes využívány zejména v oblasti hodnocení vědy. Výsledky citačních analýz jsou přímo závislé na zdroji dat, na jeho excerptci, kvalitě záznamů a jejich metadat. Při interpretaci výsledků citačních analýz je nutno vždy přihlížet k možnostem a mezím použitých metod a zdrojů.

### **Rizika a omezení citačních analýz**

Citační analýzy jsou omezovány mnoha aspekty, které mohou výsledky zkreslovat. Zásadní je *identifikace elementů*, je velmi těžké určit, zda se jedná o téhož autora, instituci, velmi složité je identifikovat úroveň elementu (výzkumný tým) a vzájemnou spolupráci, obdobné problémy jsou s identifikací publikace. Identifikaci z pohledu textu pak znesnadňují homonyma, synonyma, zkratky a různé variantní názvy. Velmi složité je určit typologii dokumentů.

Mezi technická omezení řadíme zejména otázku vzájemného propojování citací (vychází z nedokonalé identifikace). Citační analýzy předpokládají citační morálku, třetím omezením je morální aspekt. Na etické rovině hovoříme zejména o

---

\* v angl. orig. visibility

nadměrném citování sebe sama (autocitace), vzájemné bezdůvodné citování a zneužívání spoluautorství (uvedení autorů, kteří na publikaci nemají žádný či jen minimální podíl).

V jednotlivých oborech se výrazně liší citační zvyklosti; v molekulární biologii a biomedicíně je snazší dosáhnout citovanosti než v teoretické fyzice (totéž se do určité míry projevuje i uvnitř oborů). Vyšší citovanost přináší bádání na pomezí oborů, v interdisciplinárních oblastech.

Na obecnější úrovni si pak musíme uvědomit, že často mnoho pramenů (neformálních) citováno není vůbec (konzultace s kolegy aj.), že stále jsme omezeni daným zdrojem dat a nikdy se nemůžeme vyvarovat lidské potažmo úřední chybě.

Mimořádně vysokou citovanost také mívají přehledové články a některé metodické práce. Ačkoli vysoká citovanost pravděpodobně indikuje, že daná práce či její autor přináší pro odbornou komunitu něco důležitého a objevného, je třeba mít stále na paměti, že pro vědu mají velký význam i lidé, jejichž originální badatelský přínos není velký, ale jsou skvělými popularizátory, mají organizační schopnosti nebo umějí dobře přednášet. Podrobněji jsou omezení a rizika rozebrána v kapitole Scientometrické a bibliometrické indikátory.

### **Citační rejstříky**

V dnešní době se nástroje citačních analýz díky výpočetní technice rychle rozvíjejí. Klasickým nástrojem je *citační rejstřík* [SOUČEK, 2006] (též citační index), který má za cíl mapování vědy, prestiže autorů, časopisů, oborů, pracovišť atd. Otcem citačních rejstříků je Eugene Garfield, který definoval citační rejstřík, pro jejich vývoj založil Institute of Scientific Information ve Philadelphii r. 1958. Citační rejstřík je ve své minimální podobě seznam publikovaných článků opatřený výčtem jejich citací v daném časovém období. Základní metodou je tedy sledování a zaznamenávání citací a jejich počtu a slouží tak k průzkumu citovanosti, respektive jsou zdrojem pro vytváření citačních analýz. V dnešním kontextu můžeme chápat citační rejstřík šířeji, v podstatě lze jakýkoliv informační zdroj, který disponuje údaji o vzájemném propojení dokumentů pomocí citací či referencí, označit za citační rejstřík. Nebo můžeme koncept uchopit obráceně a říci, že citační rejstříky jsou dnes již automatickou součástí tzv. citačních databází, které nabízí další nástroje k citačním analýzám.

Citační rejstříky lze rozdělit na několik druhů [SOUČEK, 2006]:

- *polytematické* - zahrnují více oblastí vědy, nejznámější Science Citation Index (v rámci databáze ISI Web od Science na platformě ISI Web of Knowledge), novější databáze Scopus (<http://www.scopus.com>), Google Scholar (<http://scholar.google.com>).
- *oborové*
  - o v rámci databáze Medline (<http://www.medline.com>)
  - o astronomie, fyzika - NASA ADS (<http://adsabs.harvard.edu>)
  - o medicínský citační index Národní lékařské knihovny (<http://www.nlk.cz>)
  - o preprintová digitální knihovna pro fyziku v Los Alamos (<http://www.lanl.gov>)
  - o a další...
- *specializované* – jsou orientovány na specifickou formu vědecké komunikace, např. patentový citační rejstřík Derwent Innovation Index, STN International aj.

Nejznámější (a původní) citační rejstříky jsou:

- SCI = *Science Citation Index* (od r. 1961) – ukazatel citací v přírodních, lékařských a technických vědách
- SSCI = *Social Science Citation Index* – zaměřený na společenské vědy (od r. 1969)
- AHCI = *Arts and Humanities Citation Index* (od r. 1978) – obsahuje data z oblasti humanitních věd a umění

Jsou dostupné v rámci databáze ISI Web of Science, analytické výsledky z těchto dat jsou podrobně vypracovávány v databázi Journal Citation Reports.

Podrobnější informace k citačním rejstříkům, databázím, které substituují citační rejstříky a dalším zdrojům pro citační analýzy jsou uvedeny v kapitole Citační databáze a scientometrické zdroje.

## 4 SCIENTOMETRICKÉ A BIBLIOMETRICKÉ INDIKÁTORY A JINÉ METRIKY

*„Každý indikátor má své přednosti a svá omezení. Nesmí se na ně hledět jako na „absolutní“ ukazatele; mají komplementární roli. Různé bibliometrické postupy a metody musí být používány v kombinaci, přestože mohou někdy vést k protichůdným výsledkům, zvláště pokud poskytují užitečné informace a odpovídají vědeckým a odborným standardům. Přes svá omezení, bibliometrické ukazatele poskytují zásadní kvantitativní měřítka vědeckého výkonu.“*

[OKUBO, 1997]

Scientometrie je věda, která se snaží vstupy a výstupy ve vědě uchopit, zejména kvantitativně. K tomu potřebujeme definované metriky, na základě nichž můžeme právě kvantitu určit; určujeme ji indikátory, v kontextu evaluativní scientometrie hovoříme o scientometrických prvcích. V průběhu vývoje scientometrie bylo navrženo velké množství indikátorů, které charakterizují scientometrické elementy, neboli určují jejich status [BOLLEN, RODRIGUEZ, SOMPEL, 2008]; a dnes jsou velkým tempem navrhovány a vyvíjeny indikátory další. Nejtradičnějším indikátorem je impakt faktor, navržený Eugenem Garfieldem a další z kolekce Institute for Scientific Information (Citing Half-life aj.), dnes mu adekvátní konkurenci tvoří H-Index. Jelikož všechny indikátory mají svá omezení, nevýhody a zkrslují svým způsobem výsledky, vždy jsou navrhovány indikátory komplementární, které se snaží tato omezení zmenšit a redukovat zavádějící výpovědní hodnoty.

Pokud hovoříme o indikátorech a jejich definicích, velmi jasně vstupujeme do sporu tvrdých věd a věd společenských (sociální, humanitní). Na společenské vědy lze indikátory aplikovat jen v omezeném měřítku, neboť společenské vědy mají velmi odlišné charakteristiky vědní komunikace, ať už se jedná o typ publikací, poločas rozpadu nebo regionální vymezenost. Proto zejména ve společenských vědách je význam indikátorů velmi relativizován a je stále zdůrazňován význam recenzního řízení. Rozdíly mezi hodnotami indikátorů jsou samozřejmě i mezi jednotlivými vědeckými disciplínami, nelze jednotlivé obory, potažmo tak ani vědce, či instituce, srovnávat na základě absolutní hodnoty indikátoru. Obecně jsou však indikátory určeny pro hodnocení na všech úrovních, na mikro- (autor, článek), střední (výzkumný tým, projekt, instituce) i makroúrovni (velké instituce, stát, region).

Jak je zmíněno na začátku, indikátory hovoří o kvantitativních aspektech. V dnešní době mnoho systémů hodnocení vědy jej přebírají jako náhradní evaluační nástroj. Právě vzhledem k mnoha omezením, které jednotlivé indikátory mají, nelze nahradit kvalitativní posuzování, stejně jako nelze brát výstupní hodnoty jednotlivých metrik jako absolutní bernou minci. K rozhodnutí o zapojení indikátoru jako hodnotícího prvku je nejprve nutné se důsledně seznámit s jeho vlastnostmi, definicemi a užitými metodami tak, aby alespoň co nejlépe odpovídal záměru hodnocení.

### 4.1 OMEZENÍ INDIKÁTORŮ

#### Identifikace měřených elementů

Zásadní problematikou při hodnocení elementů je jejich identifikace, určení, jejich vymezení. Omezení lze rozdělit na technická, obsahová a etická, vždy narážíme na zásadní problém kategorizace a klasifikace na všech úrovních či právě nedodržení vědecké morálky. V případě omezení v kombinaci obsahové i technické úrovně jmenujme například identifikaci autora samotného. Pro vymezení autora nepostačuje jméno, je nutné ho určit dalšími metainformacemi, jako např. obor, instituce. Bohužel však při trendu dnešní vědy k akademické mobilitě a spolupráci je instituce naopak jako určující prvek velmi problematický. Z obsahového hlediska se jako problémové jeví vymezení disciplíny, důvody jsou tendence k interdisciplinaritě, nejasné vymezení dokumentů, zahlušení polovědeckými typy dokumentů, mnohočetné publikování, nemožnost nalézt všechny citace aj. Z technického hlediska je největším problémem propojování záznamů (citace na citovaný dokument a vice versa).

#### Datová základna a excerptce

Indikátory jsou vždy aplikovány na určitou datovou základnu, na kolekci dat, která je ovlivněna jejím výběrem. Ačkoliv pravidla excerptce se snaží držet jistou kvalitativně kvantitativní linii, nikdy není možné zajistit naprosto vyváženou sbírku. Pokud bychom uvažovali všechny výsledky, nemusí být zaručena kvalita zvažovaných výstupů a do výsledků by tak mohli zasáhnout např. chronicky píšící autoři; nehledě na to, že obsáhnutí veškeré vědecké produkce je s velkou pravděpodobností utopie.

#### Charakteristiky vědních disciplín

Jak již bylo zmíněno, každá vědecká disciplína má jiné charakteristiky, lze říci, že každou vědní disciplínu vykazuje jisté charakteristiky vědní komunikace; liší se z hlediska času, četnosti, typu publikací aj.

### **Přehledové výstupy**

Zde se dostáváme k problému jakési užitečnosti či užítosti. Review neboli přehled je obvykle sepsán vědcem, který se v oboru již dobře orientuje a je schopen uchopit problematiku a systematicky ji popsat, zároveň jméno autora často přináší jistou garanci obsahu. Takový článek je pro ostatní velmi cenný z hlediska užítku, neboť „zkracuje“ pro ostatní cestu *ad fontes*, tedy k původním pramenům, které by jinak museli složitě objevovat v původních vědeckých výstupech. Z toho plyne, že článek dosahuje vysoké citovanosti, aniž by však měl samotný vědecký přínos, aby předkládal vědecké komunitě cokoliv nového. Navíc se v tomto případě nejedná jen o články, ale samozřejmě knihy, skripta aj. Stejný problém lze definovat i mezi klinickým a základním výzkumem, kdy jeden nutně používá definice druhého, zpětně však tato vazba neexistuje.

### **Autocitace**

Autocitace je velmi známý a diskutovaný etický problém, neboť vědec by v seznamu použité literatury neměl nadbytečně odkazovat sám na sebe. Uvádí se, že maximální únosná míra autocitace je 20% [JANOŠ, 2001]. V současné době, kdy je mnohem jednodušší a efektivnější kontrola autocitací, se odborná veřejnost na základě rozsáhlých analýz začíná klonit k názoru, že autocitace nejsou zásadním problémem ve výsledcích, že je ovlivňují jen málo a spíše kolem této problematiky panuje „dlouhodobá hysterie“. Některé nástroje dnes již umožňují v předdefinovaných analytických nástrojích autocitace vyloučit. Při vylučování takových citací však musí být zvážen fakt, že autor se zpravidla systematicky věnuje jedné oblasti, jeho odkazování na starší díla je pak zcela logické a adekvátní.

### **Časová perioda**

Pokud indikátor není sám definován v rámci času (Impact Factor), podstatně do jeho výsledku právě volba časové periody zasahuje. U zkoumaných dat v dané disciplíně musíme vždy uvážit potřebný výsledek a podle toho volit, v případě multioborových analýz se opět potýkáme s problematikou rozdílů v charakteristice vědní komunikace. Každý indikátor je časem zkreslován jiným způsobem.

### **Spoluautorství**

Spoluautorství prvořadě nelze označit za pouze etické omezení. Chápáno v kontextu etiky se jedná pouze o praxi uplatňovanou zejména v některých oborech, např. v medicíně, kdy se autorská odpovědnost přisuzuje nejen opravdovému autoru, nýbrž i jeho spolupracovníkům, nadřízeným a osobám jinak spjatým s výzkumným prostředím. Druhou stránkou spoluautorství, které limituje výpovědní hodnoty indikátorů, je jeho kvantitativní uchopení. Těžko lze srovnávat velké množství výstupů vědce, který se na mnoha citovaných článcích podílel, než menší množství článků jednoho autora; nelze objektivně jasně definovat, který má větší přínos pro vědu. Počet spoluautorů je jedním ze znaků komunikace jednotlivých vědních disciplín.

### **Nesrovnatelnost napříč vědními disciplínami**

K této problematice již bylo dost zmíněno – v anglickém pojmosloví se často označuje jako *citation pattern*. Jedná se o rozdílnost vědní komunikace v jednotlivých disciplínách a některých dalších charakteristik, jako je zastarávání informací, i když pokud chápeme komunikaci v tom nejširším kontextu, je možné i zastarávání informací sem zařadit. Mezi typické komunikační charakteristiky patří počet časopisů, počet vydávaných článků, průměrný počet citací na článek.

### **Negativní citace**

Byť velmi omezeným jevem, ale přeci jen zkreslujícím jsou takzvané negativní citace. Jedná se o situaci, kdy vědec publikuje dílo, které je nekvalitní nebo má jiné závažné nedostatky, ve vědecké komunitě to vyvolá patřičnou reakci a článek pak vykazuje velkou citovanost, neboť na něj každý poukazuje. Tento jev nelze žádnou automatickou metodou vyloučit, zde si je nutné uvědomit, že právě v tuto chvíli velmi závisí na povědomí a znalostech hodnotitele. Nicméně je prokázáno, že negativní citace výsledky zásadním způsobem nezkracují. Navíc negativní citovanost často se projevuje jen v krátkém časovém intervalu.

### **Open Access**

Aspektem, který se do hodnot indikátorů může projevovat, je míra přístupu k informacím. Jestliže je publikace volně dostupná, hodnota indikátoru může stoupat [OpCit, 2008], neboť publikace má větší možnosti získat své publikum a následně citace.

## 4.2 IMPAKT FAKTOR

„Dobrý sluha, ale zlý pán“ [ŠPÁLA, 2006]

O impakt faktoru bylo napsáno mnohé, jeho dlouholetá historie ho podrobila řadě kritiky, která byla impulzem pro vznik nových indikátorů, nicméně impakt faktor je dodnes velmi uznávaným indikátorem, mnoho systémů hodnocení vědy využívá tuto metriku jako součást evaluace výstupů z vědy.

**impakt faktor = počet citací / počet publikací**

*Impakt faktor vyjadřuje impakt individuálních časopisů, podle toho, jak byly průměrně citovány ve 2 předchozích letech (v roce 2004 je IF určen za r. 2002-03). Impakt faktor je uváděn na tři desetinná místa.*

Celý název tohoto indikátoru je Journal Impact Factor\*, který avizuje, že je primárně určen pro aplikaci na časopisy, nikoliv na jiné scientometrické elementy, nelze jím hodnotit autory nebo články samotné. Při takovéto aplikaci se dopouštíme chyby, kdy uvažujeme, že úspěšný článek mnohokrát citovaný je totožný s tím, který takřka nebyl citován, pouze byl uveřejněn ve stejném časopise, potažmo autoři jsou v tuto chvíli posuzováni jako stejně úspěšní, mají stejný impakt faktor [EGGHE, 2007].

Impakt faktor je přidělován časopisům, které jsou excerpovány do tzv. citačních rejstříků, v současné době tedy do databáze ISI Web of Science, jejímž producentem je korporace Thomson Reuters. Databáze zahrnuje citační rejstříky *Science Citation Index*, *Social Science Citation Index* a *Arts & Humanities Citation Index*. Databáze ISI Web of Science je zpřístupněna mj. v integrovaném rozhraní *ISI Web of Knowledge*, které zahrnuje také databázi *Journal Citation Reports*, v níž jsou publikovány právě hodnoty impakt faktorů a další indikátory derivované z citačních rejstříků.

Z impakt faktoru lze vyvozovat, s jakou pravděpodobností lze očekávat, že články v časopise uveřejněné by mohly být citovány [ŠPÁLA, 2006]. Neboť právě časopis, ve kterém byl vydán, má přidělen takový impakt faktor, kolik v průměru jeho články získávají citací, lze pak předpokládat, že článek by tento počet citací mohl získat taktéž.

### **Historie**

Impakt faktor byl definován v r. 1955 v rámci projektu citačních rejstříků Eugenem Garfieldem, který pro zavedení tohoto indikátoru založil Institute for Scientific Information ve Philadelphii (USA). V r. 1960 vzniká první obecný citační rejstřík, o rok později byl zpracován první *Science Citation Index*. Následovalo vydání *Journal Citation Reports*, dnes je vše dostupné v rámci integrovaného rozhraní producenta Thomson Reuters ISI Web of Knowledge.

### **Kontroverzní impakt faktor:**

Sám Eugene Garfield se ve svých posledních člancích věnuje fenoménu impakt faktoru a jeho kritice, respektive jeho použití. Vyskytují se pochyby o uvádění impakt faktoru na 3 desetinná místa. Pomocí impakt faktoru nelze srovnávat časopisy z různých vědeckých disciplín nebo hodnotit autory, impakt faktor se liší podle časové periody, ne ve všech případech při delší časové periodě impakt faktor stoupá [ŠPÁLA, 2006].

Nejrozsáhlejší kritiku konceptu impakt faktoru patrně provedl v r. 1997 Per O. Seglen [SEGLEN, 1997] [HJØRLAND, 2007]. Zásadními body jsou:

- není statisticky reprezentativní pro individuální články v časopise
- nekoreluje s aktuálními citacemi jednotlivých článků
- autoři i vybírají podle mnoha jiných kritérií, když nabízejí článek do časopisu
- citace necitovatelných<sup>†</sup> jednotek jsou v databázi započítávány
- nereflektuje nadměrnou citovanost publikací typu review

---

\* Napříč učebním textem je dle zvyklostí v českých odborných textech používán název „impakt faktor“.

<sup>†</sup> v originále non-citable

- rozsáhlé články získávají více citací, hodnota impakt faktoru se zvyšuje
- navyšování citací pomocí autocitací u časopisů s kratší periodou vydávání, jinak též selektivní autocitace, články mají tendenci přednostně citovat jiné články ve stejném časopise
- autory časopisu jsou preferovány citace v národním jazyce časopisu
- omezené pokrytí databáze
- v databázi nejsou zahrnuty knihy, i když jsou významným zdrojem citací
- orientace databáze zejména na anglický jazyk, dominance amerických publikací
- kolekce časopisů v databázi se může rok od roku lišit
- vyšší hodnoty mají výzkumné oblasti s rychle zastarávajícími informacemi
- závislost na dynamice oboru v daném čase
- menší vědecké disciplíny většinou nemají časopisy s vysokým impaktem
- je silně ovlivněn vztahy mezi jednotlivými disciplínami (např. klinický vs. základní výzkum)
- je určen mírou citovanosti, což neplatí vice versa
- ignoruje kvalitu přijatých citací, rozpor mezi prestiží a popularitou

#### 4.3 DALŠÍ KLASICKÉ BIBLIOMETRICKÉ INDIKÁTORY INSTITUTE OF SCIENTIFIC INFORMATION

Spolu se zavedením impakt faktoru byly zavedeny další indikátory, jejichž hodnoty jsou určovány z datové kolekce citačních rejstříků. Indexy jsou určován buď pro jednotlivé časopisy, nebo v rámci předmětné oblasti, tzv. agregované indexy.

##### **Immediacy Index**

Immediacy Index, v češtině nazýván jako index bezprostředního vlivu nebo odezvy, někdy označován také jako Garfieldův index, vyjadřuje průměrný počet citací, který článek získal v roce jeho publikování.

**immediacy index = počet citací článku v daném roce / počet článků v daném roce**

Vzhledem k tomu, že je uvažován průměr citací na článek, immediacy index redukuje vliv velkých časopisů. Časopisy s velkou periodicitou jsou zvýhodněny, neboť mají více času v průběhu roku nasbírat citace. Časopisy, které vycházejí nepravidelně nebo ke konci roku tak mají nižší immediacy index. Tento index je vhodný zejména pro srovnávání časopisů zabývajících se velmi aktuálními otázkami a dynamickými vědními disciplínami.

##### **Cited Half-life**

*Poločas citovanosti*, jak je do češtiny tento ukazatel překládán, udává po kolika letech (ve vztahu k určitému roku) se objeví 50% všech citací na články daného časopisu v citačních rejstřících. Například časopis, který v roce 2001 má cited half-life 5, což znamená, že články publikované mezi 1997-2001 (včetně) odpovídají polovině všech citací článků z tohoto časopisu v roce 2001. Cited half-life je vypočítáván pouze pro ty časopisy, které byly citovány více než stokrát v daném roce.

Cited half-life má výpovědní hodnotu zejména pro akvizici a archivování, změny mohou naznačovat změnu ve formátu časopisu a dá se z nich vyvozovat historie časopisu.

##### **Citing Half-life**

*Poločas citování* vyjadřuje průměrné stáří článků, které jsou citovány v daném časopise v jednom roce. Např. pokud má časopis v r. 2008 hodnotu citing half-life 9, značí to, že polovina článků, které byly články v průběhu roku v daném časopise citovány, byly publikovány mezi lety 2000-2008 (včetně). Obdobně jako cited half-life je tento index vypočten jen u časopisů s více jak 100 referencemi.



## 4.4 H-INDEX

*H-index* je poměrně nový indikátor, byl definován v roce 2005 fyzikem Jorge Hirschem. tento indikátor byl vědeckou komunitou velmi rychle akceptován, dnes je již plně respektovaným ukazatelem. Název H-index byl původně odvozen právě od příjmení autora Hirsch, dnes se však již využívá spíše jeho charakteristiky a  $h$  je interpretováno jako *highly cited index*.

H-index je číslo  $h$ , které **udává počet publikací, které byly alespoň  $h$ -krát citovány**. Pokud tedy vědec má h-index roven 5, značí to, že nehladě na to, kolik publikací celkově napsal, 5 jeho publikací bylo nejméně 5krát citováno.

H-index lze jednoduše odečíst i ze seznamu publikací seřazených podle počtu citací, kdy h-index je průsečíkem pořadí dokumentu a udaného počtu citací.

Př.:

pořadí	název	počet citací
1	Dokument A	20
2	Dokument B	8
3	Dokument C	5
<b>4</b>	<b>Dokument D</b>	<b>4</b>
5	Dokument E	1

**$h = 4$ : pořadí = počet citací**

H-index může být aplikován na jakýkoliv scientometrický element na mikroúrovni (pro hodnocení autorů byl tento indikátor původně vytvořen, jinak např. časopis) a na střední úrovni (výzkumný tým, instituce aj.). Pomocí H-indexu lze odvozovat také vývoj vědecké disciplíny, pokud provedeme rešerši na tuto oblast a seřadíme dle citací, definovat aktuální témata a naopak eliminovat témata zastarávající [BORNMANN, 2007].

### Výhody H-indexu

Fakt, že byl H-index odbornou komunitou velmi rychle přijat, svědčí o tom, že má mnoho výhod. Je především určen pro efektivní hodnocení na mikroúrovni, pokrývá jak množství tak „viditelnost“ [BORNMANN, 2007], tedy neuvažuje málo citované publikace nehladě na jejich množství [EGGHE, 2007] a minimalizuje vliv vědců s jednou citovanou publikací, kteří však jiný přínos nemají (typicky autor publikace typu review). Hlavní předností je též jeho jednoduchost a odvoditelnost z jednoduchého seznamu [BORNMANN, 2007].

### Nevýhody H-indexu

To, co je na jedné straně interpretováno jako výhoda, se může z obráceného úhlu pohledu jevit též jako nevýhoda. Konkrétně h-index nebere v potaz, je-li článek dále citován a jestli dále obdrží 100, 1000 nebo více citací [EGGHE, 2007], tj. nerozlišuje již dále kvalitní vědce, kterých je v populárních disciplínách mnoho. Indexu je taktéž někdy označován za index průměrnosti, neboť výrazně zvýhodňuje vědce dlouhodobě působící v oboru, nicméně v takovém případě lze velmi jednoduše aplikovat časové omezení a naopak umožňuje srovnání vědců různého věku [ŠPÁLA, 2006]. Právě možnost určení časové periody pomocí indexu lze sestavit jakousi klasifikační stupnici. Např. pokud vědecký pracovník po 20 letech práce dosáhne  $h$  rovné 20, je ho možno označit za úspěšného. Pokud  $h$  bude rovné 40 je vynikající a při hodnotě 60 (nebo 90 po 30 letech) je skutečně výjimečným jedincem. [ŠPÁLA, 2006].

Taktéž přílišná jednoduchost se stává terčem kritiky, otázkou zůstává, zda lze měřit vše jen jedním indexem a redukovat tak vícestranný pohled [GLÄNZEL, 2006]. H-index dodnes taktéž není ověřen ve všech vědních oblastech [BORNMANN, 2007], jeho funkčnost je dokázána zejména v některých tvrdých vědách (fyzika, biologie, ekologie, evoluční biologie). Opět se vracíme k obecnému omezení většiny indikátorů, kdy h-indexem taktéž lze porovnávat pouze vědce působící v jedné vědní oblasti, neřeší problém rozlišných charakteristik vědních oborů. Sám Hirsch konstatuje, že index neodstraňuje problém review, které vždy získává více citací než článek s původním výzkumem [ŠPÁLA, 2006].

---

\* anglický termín *visibility*



## **Revize h-indexu a nové indikátory z něj odvozené**

Všechna omezení h-indexu vedla k zamyšlení a pokusu limitování daného parametru, který do této metriky vstupuje, a dalšímu zdokonalení tohoto indikátoru. Jak již bylo zmíněno, ihned vešlo v používání omezení na danou časovou periodu. Problematika autocitací lze taktéž poměrně jednoduše řešit tak, že publikace budou do výpočtu započteny jen v tom případě, kdy po odečtení autocitací převyšuje výsledný počet citací hodnotu h-indexu. Bornmann [2007] a Hirsch se shodli taktéž na řešení problému rozdílu mezi jednotlivými disciplínami na základě jejich charakteristiky spoluautorství; pokud sečteme všechny h-autory (multiplicitně) a podělíme h-indexem, dostaneme se k hodnotám, které rozdíl mezi jednotlivými disciplínami redukuje a můžeme je tak srovnávat [BATISTA, 2006].

K h-indexu byly definovány další komplementární indikátory, které vychází z jeho podstaty, jako např. a-index nebo g-index (viz následující podkapitola o dalších indikátorech).

## **4.5 DALŠÍ INDIKÁTORY**

### **4.5.1 G-index**

Specialista na Lotkuv zákon Leo Egghe inspirován nedostatkem h-indexu, který nereflektuje rozlišení vědců dle výkonnosti jejich nejlepších článků, navrhl nový indikátor, *G-Index*, který umožňuje rozlišit v této úzké skupině výkonnější autory. G-index je definován jako g počet publikací, které dohromady získaly  $g^2$  a více citací. Jinými slovy, čím vyšší počet citací nejlepších (nejcitovanějších) článků, tím vyšší bude hodnota g-indexu. G-index vždy převyšuje skóre h-indexu a ukazuje výkonnost nejlepších článků, respektuje tak nejvíce citované články [EGGHE, 2007] [BORNMANN, 2007].

### **4.5.2 Eigenfactor**

Jednu z nových metod určených k hodnocení vědy nabízí portál Eigenfactor.org s podtitulem „Ranking and Mapping Scientific Knowledge“. Pro hodnocení portál používá dva indikátory *Eigenfactor Score* a *Article Influence*. Eigenfactor Score vyjadřuje míru důležitosti časopisu pro vědeckou komunitu. Podle definice a matematického modelu sítě dokumentů navzájem propojené citacemi, ve které se vědec pohybuje, je to odhad procenta času, který stráví uživatelé s daným časopisem. Také proto se toto skóre zvětšuje s velikostí časopisu, je přímo závislé na počtu článků, větší časopisy mají také větší Eigenfactor Score. Skóre je uváděno jako percentil, tj. součet skóre všech časopisů je roven 100. V r. 2006 měl nejvyšší Eigenfactor Score časopis Nature, a to 2.138; v prvních tisícovce časopisů skóre neklesá pod 0.01.

Druhý indikátor, Article Influence, je určen k odhadnutí důležitosti časopisu neohledně na jeho velikost (počet publikovaných článků). Article Influence vyjadřuje míru průměrného vlivu každého článku v časopise za pět let po jeho publikování. Pokud je jeho hodnota u časopisu např. 20, znamená to, že průměrný článek v tomto časopise je dvacetkrát vlivnější než průměrný článek v celé kolekci dat. Tento indikátor je srovnatelný s impakt faktorem.

### **4.5.3 Y-factor**

*Y-factor* je indikátor navržený mj. Herbertem Van de Sompelem, který již stál za zrodem velmi významných věcí jako např. OpenURL [BOLLEN, RODRIGUEZ, SOMPEL, 2008]. Tento indikátor na rozdíl od předchozích, které se profilují jako komplementární k ostatním indikátorům, se snaží být syntetický, princip spočívá v řazení časopisů na základě jejich impakt faktoru a vážené hodnoty PageRanku (Weighted PageRank, [XING, GHORBANI, 2004]).

$$Y = IF \times PR_w$$

Y-faktor zavádí do hodnocení časopisů pojem popularity a prestiže. Vzhledem k tomu, že impakt faktor je závislý na četnosti citací, vyzdvihuje tak spíše aspekt popularity časopisu. Weighted PageRank svou podstatou závisí více vzájemném respektu mezi prestižními časopisy, je jim přisouzena vyšší hodnota, pokud citace pochází z vlivnějšího časopisu a lépe odráží náš podvědomý koncept prestiže; jinak řečeno, neohlídá jen na kvantitu citací, ale též na jejich kvalitu.

Populární časopisy jsou ty, které jsou často citovány časopisy s nízkou prestiží. Tyto časopisy mají velmi vysoký impakt faktor a velmi nízký Weighted PageRank. Prestižní časopisy jsou ty, které nejsou tak často citovány, ale jejich citace pocházejí z vysoce prestižních časopisů. Tyto časopisy mají nízký impakt faktor a vysoký Weighted PageRank.

#### 4.5.4 A-index / Jinův index

Tento index je komplementární k h-indexu, kdy Jin a Egghe charakterizují necitlivost h-indexu k jednomu nebo více vysoce citovaným dokumentům jako nevýhodu, tato míra by měla respektovat nejcitovanější články. A-index udává průměrný počet citací v h dokumentech [BORNMANN

#### 4.5.5 Web Impact Factor

Web Impact Factor se řadí mezi kvantitativní indikátory, který měří průměrné frekvence odkazů, je obrazem obsahu pokrytí vyhledávacího nástroje v určitém čase. [NORUZI, 2006]. Autorem je Peter Ingwersen, indikátor definoval v r. 1998 a dodnes jsou o něm vedeny diskuze, doporučuje se používat jej opatrně. U tohoto indikátoru vyvstává problematika spojená typicky v databázích s formou excerptce, transformovaná v prostředí internetu na validitu obsahu, kdy webový zdroj není recenzován a tudíž neprochází kontrolou kvality. Web Impact Factor, ačkoliv je inspirován klasickým impakt faktorem, není ekvivalentní. Web Impact Factor vyjadřuje ve své podstatě „atraktivitu“ webových stránek nebo domén na webu.

#### 4.5.6 Citační / referenční denzita

Citační a referenční denzita je jedna s klasických charakteristik každé vědní disciplíny. Vyjadřuje průměrné množství citací / referencí na článek.

#### 4.5.7 Erdösovo číslo

Indikátor, který se zcela odlišuje od všech předchozích koncepcí je tzv. Erdösovo číslo, jako jeden z mála obsahuje výrazný sociální aspekt. Erdösovo číslo je určeno pouze v matematice, indikuje topologickou vzdálenost v grafu znázorňujícím vztahy spoluautorů. Paulu Erdösovi je přisouzena hodnota 0, autoři, kteří se podíleli na některém z jeho děl, mají hodnotu 1, jejich spoluautoři, kteří však již nejsou spoluautory Erdöse 2 atd.

Uvedme, že P. Erdős byl maďarsko-americký matematik (1913-1996) a velmi produktivní autor (více než 1.000 publikací). E je celkem zvláštní překvapivá míra výkonnosti, známá přes více než 25 let. Číslo E je fenomén vlastní matematické vědecké komunitě z druhé poloviny minulého století, kde panovala silná spolupráce s průvodním spoluautorstvím. Nositelé různých ocenění, např. Fieldsovy medaile dosahovali koncem minulého století E nižší než 6 a více než 60 nositelů Nobelovy ceny za fyziku, chemii, ekonomiku a medicínu mají E nižší než 9. Jsou sestaveny rozsáhlé seznamy hodnoty E u nositelů různých cen a medailí (např. „Fields“, „Nevalinna“, „Wolf“, „Steel“) jakož početné graficky vyjádřené cesty počínající vždy Erdösem a sestupně se odvíjející podle následných spoluautorství. Pro ostatní obory E asi nemá význam, i když u některých autorů (např. z medicíny F. H. C. Cricka, genetika a biofyzika, spolunositel Nobelovy ceny za objev dvoušroubovice DNA), E dosáhlo hodnoty 7 [ŠPÁLA, 2006].

#### 4.5.8 Euro-factor

Hofbauer a jeho kolegové v r. 2002 navrhli novou metriku pro evropské časopisy, reagující tak na to, že jediná uznávaná metrika je 40 let stará a je americké provenience s nedostatečnou excerptcí evropských zdrojů. Tento indikátor se neujal.

## 5 CITAČNÍ DATABÁZE A SCIENTOMETRICKÉ ZDROJE

Jednou z kvalifikací informační profese je znalost, kde najít patřičné informace, jaké zdroje pro získání relevantních a pertinentních informací využít. Pokud informační profesionál nedisponuje těmito znalostmi, musí znát alespoň metody, jak nejrelevantnější zdroje nalézt. Jednou z možných metod je užití citačních databází, které uživatele pomocí citační a publikační analýzy nasměrují k takovým zdrojům, které již prošly jakýmsi hodnocením. Podle míry citovanosti článku (časopisu, autora) lze usuzovat na kvalitu zdroje. Toto hledisko výběru zdrojů však není vždy relevantní, je třeba si uvědomit, že kritérium míry citovanosti nemusí být vždy adekvátní, často mohou být lepším zdrojem oborové databáze.

Oborové databáze mají výhodu v tom, že zachází se specifickými informacemi. Mívají kvalitnější relevantní popisný aparát, neboť mohou zohledňovat typickou strukturu dokumentů. Mají vlastní oborové tezaury, předmětová hesla či klíčová slova. Vynikají hloubkou obsahu, jsou schopné dát odpověď na velmi specifické a přesně definované dotazy. Vzhledem k jejich orientaci na obsah však nedisponují pokročilými funkcemi jako je práce s autory či institucemi, málokdy nabízejí specifické funkce, které usnadňují práci uživatelům.

Oproti tomu *citační databáze* nabízí rychlou identifikaci významných zdrojů, publikací a autorů pro daný okruh. Jsou obsahově méně hluboké než oborové databáze, svým polytematickým zaměřením však doplňují oborové databáze o aspekt interdisciplinarity. Pro uživatele jsou k dispozici další analytické nástroje, databáze se snaží autoritně zpracovávat autory a instituce. Díky těmto nástrojům pak můžeme porovnávat časopisy, autory, zkoumat jejich vzájemnou spolupráci atd.

Pokud se zabýváme scientometrií, narazíme na několik typů zdrojů informací. Klasickým jsou již zmíněné citační databáze, do kterých řadíme ISI Web of Science s jejich citačními rejstříky, databázi Scopus, Google Scholar a některé oborové databáze, které propojují jednotlivé dokumenty na základě citací a referencí (PubMed). Tyto zdroje užíváme k analýzám výstupů ve vědě a jednomu z hodnocení vědeckých výstupů.

Především společnost Thomson Reuters rozvinula zdroje pro scientometrii v oblasti vědy o vědě a jejího hodnocení. Nabízí několik databází, které slouží k hodnocení vědy na úrovni časopisů, institucí či států. Scopus postupně dohání tyto databáze implementací funkcí přímo do databáze, které umožňují srovnatelné analýzy.

Ve scientometrii lze využít dalších syntetických nástrojů, jako jsou portály informací o nejvíce citovaných vědcích či časopisech, nebo portály, které na základě propojení dokumentů citacemi se snaží mapovat vědu a vytvářet další analýzy. Rozdělení na různé typy není kategorické, mnoho zdrojů se překrývá a lze je zařadit do více kategorií.

## 5.1 DATABÁZE PRO HODNOCENÍ VĚDY

Databáze pro hodnocení vědy můžeme rozlišit na dva základní druhy. První, z provenience společnosti Thomson Reuters [Thomson Reuters, 2008] (sekce Scientific, oblast Research and Development), které poskytují analytická a syntetická data vycházející z ISI Web of Science a citačních rejstříků a jejich využívání. Tyto databáze jsou bohužel v rámci ČR takřka nedostupné. Druhou kategorií jsou databáze na národní úrovni, které zabezpečují proces hodnocení vědy v daném státě. Za ČR je to databáze RIV, rejstřík informací o výsledcích, který je součástí informačního systému výzkumu, vývoje a inovací v ČR. Dostupný je na <http://aplikace.isvav.cvut.cz/prepareResultForm.do>. Těmto databázím se zde věnovat nebudeme.

### **Essential Science Indicators** [Thomson Reuters, 2008]

Souhrnná databáze statistik vědecké výkonnosti a vědeckých trendů založená na počtech publikovaných článků a citačních dat z databází Thomson Reuters. Je to analytický zdroj informací, který umožňuje analyzovat výkonnost vědy na různých úrovních, identifikovat vědecké trendy, určovat nejvýkonnější státy, časopisy, instituce aj., hodnotit vědecké výstupy a jejich impakt, hodnotit autory.

### **Institutional Citation Report** [Thomson Reuters, 2008]

Tato databáze je určena k analýzám vědeckých výstupů a impaktu na institucionální úrovni. Umožňuje srovnávat, řadit a hodnotit instituce.

### **National Citation Report** [Thomson Reuters, 2008]

Je obdobou výše uvedené databáze Institutional Citation Reports, je zaměřena na státní úroveň.

### **National Science Indicators** [Thomson Reuters, 2008]

Statistická databáze vědecké výkonnosti v rámci států. Obsahuje statistiky z více jak 180 států, jsou k dispozici data 25 let zpětně. Nabízí vizualizaci dat.

### **Journal Performance Indicators** [Thomson Reuters, 2008]

Nabízí data z více než 10000 vědeckých vysoce citovaných časopisů za více jak 25 let, poskytují za různá období pro různé časopisy a články různé citační a impaktfaktorové údaje. V porovnání s Journal Citation Reports poskytují přesnější údaje [ŠPÁLA, 2006].

### **Journal Citation Reports** [Thomson Reuters, 2008]

Tradiční databáze je přímým výstupem z dat citačních rejstříků Science Citation Index a Social Science Citation Index.

<b>Select a JCR edition and year:</b>	<b>Select an option:</b>
<input checked="" type="radio"/> JCR Science Edition <input type="text" value="2007"/>	<input checked="" type="radio"/> View a group of journals by <input type="text" value="Country/Territory"/>
<input type="radio"/> JCR Social Sciences Edition <input type="text" value="2007"/>	<input type="radio"/> Search for a specific journal <input type="radio"/> View all journals
<input type="button" value="SUBMIT"/>	

**Obr. 3 – Vstupní obrazovka do databáze Journal Citation Reports**

Journal Citation Reports poskytuje statistické informace a analytická data o časopisech, výsledky citačních analýz, určuje pro každý časopis bibliometrické indikátory impakt faktor, Immediacy Index, Cited Half-life. Po kliknutí na název časopisu jsou k dispozici další údaje o časopisu včetně výpočtů daných indikátorů. Databáze je vhodně propojena s databází Ulrich's Periodical, která zpřístupňuje detailní data o časopisech. Dostupná v ČR, samostatně nebo v rámci platformy ISI Web of Knowledge (<http://www.isiknowledge.com/>).

**Journal Summary List**
[Journal Title Changes](#)

**Journals from:** **subject categories INFORMATION SCIENCE & LIBRARY SCIENCE**

**Sorted by:**

---

**Journals 1 - 20 (of 56)**    Page 1 of 3

       *Ranking is based on your journal and sort selections.*

Mark	Rank	Abbreviated Journal Title <i>(linked to journal information)</i>	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-life
<input type="checkbox"/>	1	<a href="#">MIS QUART</a>	0276-7783	4329	5.826	0.533	30	9.4
<input type="checkbox"/>	2	<a href="#">J AM MED INFORM ASSN</a>	1067-5027	2394	3.094	0.699	93	5.2
<input type="checkbox"/>	3	<a href="#">INFORM SYST RES</a>	1047-7047	2146	2.682	0.130	23	8.3
<input type="checkbox"/>	4	<a href="#">ANNU REV INFORM SCI</a>	0066-4200	378	1.963	0.533	15	6.1
<input type="checkbox"/>	5	<a href="#">J MANAGE INFORM SYST</a>	0742-1222	1861	1.867	0.175	40	7.7

**Obr. 4 – Analytický přehled časopisů v oblasti Information Science and Librarianship podle impakt faktoru**

## 5.2 CITAČNÍ DATABÁZE

Jako citační databáze dnes bývají uváděny „dvě plus jedna“ základní platforma: ISI Web of Science, Scopus a opatrně je přidáván Google Scholar. Samozřejmě by bylo možné při dnešních technologických možnostech uvádět i další databáze, jako je např. PubMed, rozhraní databáze Medline, které sbírají informace citacích. Citační databáze však tyto záznamy

propojují v obou směrech a nabízejí další funkce pro citační a jiné analýzy. Právě proto se Google Scholar ke dvojici přiřazuje opatrně, neboť v současné chvíli spíše slouží jako vyhledávač než jako citační databáze.

### 5.2.1 ISI Web of Science a citační rejstříky

Databáze *ISI Web of Science* již byla zmíněna mnohokrát. Její historické kořeny, respektive předchůdci, sahají až do 60. let 20. století, kdy Eugene Garfield začal vytvářet citační rejstříky a založil Institute of Scientific Information (zkratka ISI - před jménem databáze). S příchodem informačních a komunikačních technologií vznikla počítačová verze, v dnešní podobě je to typicky zelené rozhraní ISI Web of Science. Dnes je producentem mediální a informační magnát *Thomson Reuters*, ještě nedávno spadala pod sekci Thomson Scientific, než došlo k fúzi s Reuters.

ISI Web of Science je databáze, která sdružuje několik dílčích informačních zdrojů, v závislosti na subskripci instituce. Pro ČR jsou v rámci ISI Web of Science dostupné všechny tři citační rejstříky, Science Citation Index, Social Science Citation Index, Arts and Humanities Citation Index a některé dílčí databáze.

V rámci integrace svých zdrojů společnost Thomson Reuters provozuje platformu *ISI Web of Knowledge* (<http://www.isiknowledge.com/>), ve které jsou sdruženy všechny předplacené zdroje. Je možné pracovat jak zároveň ve všech tak pouze v jedné z databází. V ČR jsou typicky v rámci ISI Web of Knowledge dostupné databáze ISI Web of Science, Current Contents Connect a Journal Citation Reports. Current Contents Connect je databáze obsahů a bibliografických informací odborných časopisů a knih. V rámci této databáze jsou zveřejňovány články dříve, než vyjdou v časopise.

*Science Citation Index* je nejstarším citačním rejstříkem, byl definován v r. 1961 a poprvé vydán r. 1963. Je to databáze převážně bibliografických informací (bez plných textů). Do tohoto rejstříku je excerpováno přibližně 3700 odborných časopisů, ve své rozšířené verzi *Science Citation Index Expanded* přes 6500, z oblasti zemědělství, biologie, ekologie, z oblasti inženýrství a technických věd, aplikovaných věd, medicíny a lékařských věd, fyziky a chemie. Informace jsou aktualizovány týdně.

*Social Science Citation Index* vznikl v r. 1973, obsahuje data z přibližně 2000 časopisů z 50 společenských disciplín. Některé časopisy mohou být zařazeny v obou dvou rejstřících. Tradičně nejsilnějšími disciplínami jsou právě ty, které mají časopisy v obou rejstřících. Patří mezi ně psychologie, psychiatrie a pečovatelské vědy. Dalšími výraznými disciplínami jsou ekonomie, management, politologie a sociologie.

*Arts and Humanities Index* je nejmladší, je datován do r. 1979. Tato databáze není zpracována jako dvě předchozí do Journal Citation Reports. Obsahuje přes 1000 časopisů z oblasti umění a humanitních věd, mezi ně patří např. disciplíny jako archeologie, umění a architektura, historie, jazykověda a lingvistika, literatura, filozofie, herectví, náboženství a religionistika; nechybí ani kategorie všeobecnosti. Tento rejstřík má vlastní specifický popisný aparát.

ISI Web of Knowledge<sup>SM</sup> Take the next step

All Databases | Select a Database | Web of Science | Additional Resources

Search | Cited Reference Search | Structure Search | Advanced Search | Search History | Marked List (0)

Web of Science®

Search for:

in

*Example: oil spill\* mediterranean*

AND  in

*Example: O'Brian C\* OR OBrian C\**  
Need help finding papers by an author? Use [Author Finder](#).

AND  in

*Example: Cancer\* OR Journal of Cancer Research and Clinical Oncology*

[Add Another Field >>](#)

Current Limits: [\[Hide Limits and Settings\]](#)

**Timespan:**

All Years  (updated 2008-10-11)

From  to  (default is all years)

**Citation Databases:**

Science Citation Index Expanded (SCI-EXPANDED)--1945-present

Social Sciences Citation Index (SSCI)--1980-present

Arts & Humanities Citation Index (A&HCI)--1980-present

**Chemical Databases:**

Index Chemicus (IC)--1993-present

Current Chemical Reactions (CCR-EXPANDED [back to 1840])--1986-present

**Obr. 5 – Úvodní obrazovka databáze ISI Web of Science s rozhraním jednoduchého vyhledávání**

Databáze ISI Web of Science disponuje uživatelsky přívětivým jednoduchým vyhledávacím rozhraním. Nabízí specifické možnosti vyhledávání jako je „Cited References Search“, hledání v citacích určeného autora a zdroje, a pokročilé vyhledávání. Lze si ukládat historii vyhledávání. Jakoukoliv sadu výsledků lze analyzovat ve smyslu počtu záznamů podle autora, země, typu dokumentu, instituce, časopisů atd. Tato data je možné si uložit.



292 records. Topic=(scientometrics)

Rank the records by this field:	Analyze:	Set display options:	Sort by:
Author ▲ Country/Territory Document Type Institution Name ▼	Up to <input type="text" value="500"/> records.	Show the top <input type="text" value="10"/> results. Minimum record count (threshold): <input type="text" value="3"/>	<input checked="" type="radio"/> Record count <input type="radio"/> Selected field

Analyze

Use the checkboxes below to view the records. You can choose to view those selected records, or you can exclude them ( and view the others).

**Note:** The number of records displayed may be greater than the listed Record Count if the original set contained more records than the number of records analyzed.

<input type="checkbox"/> View Records <input checked="" type="checkbox"/> Exclude Records	Field: Author	Record Count	% of 292	Bar Chart	Save Analysis Data to File
<input type="checkbox"/>	GARFIELD, E	12	4.1096 %	■	
<input type="checkbox"/>	GARG, KC	11	3.7671 %	■	
<input type="checkbox"/>	LEYDESDORFF, L	10	3.4247 %	■	
<input type="checkbox"/>	SCHUBERT, A	10	3.4247 %	■	
<input type="checkbox"/>	GLANZEL, W	8	2.7397 %	■	
<input type="checkbox"/>	BONITZ, M	6	2.0548 %	■	
<input type="checkbox"/>	COURTIAL, JP	6	2.0548 %	■	
<input type="checkbox"/>	KRETSCHMER, H	6	2.0548 %	■	
<input type="checkbox"/>	VINKLER, P	6	2.0548 %	■	
<input type="checkbox"/>	WOUTERS, P	6	2.0548 %	■	

(346 Author value(s) outside display options.)

**Obr. 6 – Analýza sady vyhledaných záznamů podle zastoupení autorů**

Portál nabízí personalizované funkce jako je sestavení vlastního seznamu časopisů, upozornění na citace nebo právě uložení hledání. Je efektivně propojen s dalšími službami s dílny Thomson Reuters jako je citační manažer EndNote Web (<https://www.myendnoteweb.com/>), do kterého lze přímo ukládat záznamy, jako i do dalších referenčních manažerů. Nabízí odkaz na portál MyResearcher's ID, který se snaží za pomoci uživatelů vytvořit standardizovaný seznam autorit. Je možné zadaný dotaz použít pro vyhledávání na webu pomocí služby Scientific WebPlus (<http://scientific.thomsonwebplus.com/>) zaměřené na vyhledávání odborných informací.

V poslední době se databáze hojně zaměřila na vylepšení identifikace autorů a institucí, pokud autora vyhledávání rozpozná jako autoritu, nabídne k němu tzv. Distinct Author Set, tj. souhrn všech údajů, které jsou mu v rámci databáze přiřazeny.



Select one or more sets and click "View Records".

Select	Set	Author Names	# of Records	Source Titles (top 5 by record count)	Publication Years	Subject Areas (top 5 by record count)
<input type="checkbox"/>	1.	Glanzel, W	90	- SCIENTOMETRICS (66) - INFORMATION PROCESSING & MANAGEMENT (4) - JOURNAL OF INFORMATION SCIENCE (4) - TRAC-TRENDS IN ANALYTICAL CHEMISTRY (3) - CZECHOSLOVAK JOURNAL OF PHYSICS (2)	1983 - 2005	- Multidisciplinary Science & Technology (81) - Physical Sciences (11) - Social Sciences (1)

Obr. 7 – Sada informací k danému autoru (Distinct Author Set)

Databáze umožňuje vytvořit tzv. Citation Report, který analyzuje výsledky. Zobrazuje graf publikací za rok a citací přijatých každý rok. Informuje o celkové citovanosti, průměrné citovanosti a h-indexu. Z výsledků analýzy je možné vyloučit autocitace. Obsahuje tabulku citací podle jednotlivých článků a roku.

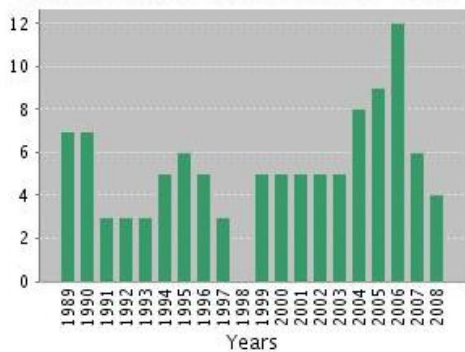
### Citation Report

Author=(GI?nzel W\*)

Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, IC, CCR-EXPANDED [back to 1840].

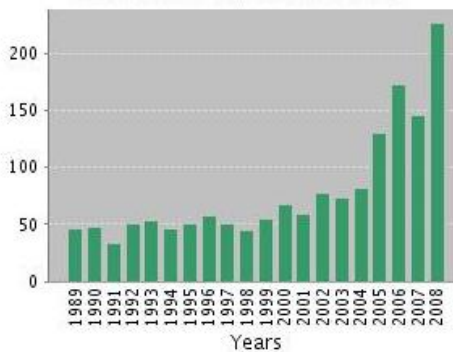
This report reflects citations to source items indexed within Web of Science. Perform a Cited Reference Search to include citations to items not indexed within Web of Science.

#### Published Items in Each Year



The latest 20 years are displayed.  
View a graph with all years.

#### Citations in Each Year



The latest 20 years are displayed.  
View a graph with all years.

Results found: 124

Sum of the Times Cited [?]: 1,629  
View Citing Articles  
View without self-citations

Average Citations per Item [?]: 13.14

h-index [?]: 23

Obr. 8 – Citační analýza článků daného autora – graf vydaných publikací v určitém roce, graf citací přijatých každý rok. Napravo další statistické údaje.

Use the checkboxes to remove individual items from this Citation Report or restrict to items processed between		2004	2005	2006	2007	2008	Total	Average Citations per Year
<input type="checkbox"/> 1. Title: SCIENTOMETRIC DATAFILES - A COMPREHENSIVE SET OF INDICATORS ON 2649 JOURNALS AND 96 COUNTRIES IN ALL MAJOR SCIENCE FIELDS AND SUBFIELDS 1981-1985 Author(s): SCHUBERT A, GLANZEL W, BRAUN T Source: SCIENTOMETRICS Volume: 16 Issue: 1-6 Pages: 3-& Published: JUN 1989		4	4	8	1	1	133	6.65
<input type="checkbox"/> 2. Title: Journal impact measures in bibliometric research Author(s): Glanzel W, Moed HF Source: SCIENTOMETRICS Volume: 53 Issue: 2 Pages: 171-193 Published: FEB 2002		3	13	15	15	12	61	8.71
		82	130	172	145	226	1,629	62.65

Obr. 9 – Analýza citací každého článku podle jednotlivých let

Databáze ISI Web of Science obsahuje snadno ovladatelné nástroje umožňující základní scientometrické analýzy. Je dnes respektovaným zdrojem informací, i když při tom, jak stabilní a dlouholetou pozici na trhu má, se samozřejmě vyskytuje řada kritických ohlasů. Co se týče uživatelského rozhraní, trochu zaostává za svou konkurencí, databází Scopus, která je podstatně intuitivnější a více zaměřená na uživatele. Velmi těžké je v databázi zjistit, jak prohlížet obsažené zdroje. Obsahově se o databázi ISI Web of Science hovoří jako o úzké a hluboké, kdežto Scopus je považován za mělký a široký. Obojí dvojí může být i výhodou i nevýhodou. Velkým přínosem je to, že existují dva silně konkurenční produkty, což tlačí vývoj jejich možností velkou rychlostí dopředu a snaží se dohánět vzájemné nevýhody.

### 5.2.2 Scopus

Scopus ([www.scopus.com](http://www.scopus.com)) je citační a abstraktová databáze produkovaná tradičním vydavatelem odborné literatury společností Elsevier (založena r. 1880). Byla spuštěna v r. 2004. Excerptuje přes 15 000 vědeckých časopisů ze všech oblastí vědy. Zahrnuje informační zdroje z největších patentových databází, z webu a dalších zdrojů. Retrospektiva databáze je do r. 1966, je aktualizována denně. Citační a jiné analýzy lze provádět s daty od r. 1996, starší záznamy nejsou plně zpracovány. V ČR není vyjma některých ústavů Akademie věd ČR a České geologické služby k dispozici.

Scopus completes extensive archive project - backfilling 2,450 journals to volume 1, issue 1. [Read more...](#)

Brought to you by The Scopus Team [Ask a Librarian](#)

The screenshot shows the Scopus search interface with the following elements:

- Navigation tabs:** Basic Search (selected), Author Search, Affiliation Search, Advanced Search.
- Search bar:** "Search for:" followed by a text input field and "in" followed by a dropdown menu set to "Article Title, Abstract, Keywords". Below the input field is the example text: "E.g., 'heart attack' AND stress".
- Advanced search options:** A second "AND" dropdown menu followed by another text input field and a dropdown menu set to "Article Title, Abstract, Keywords".
- Buttons:** "Search" and "Clear" buttons are located below the search bar and again below the advanced search options.
- Limit to:**
  - Date Range (inclusive):**
    - Published: "All years" to "Present"
    - Added to Scopus in the last "7" days
  - Document Type:** A dropdown menu set to "All".
  - Subject Areas:**
    - Life Sciences (> 3,400 titles)
    - Health Sciences (> 5,300 titles) Includes 100% Medline coverage
    - Physical Sciences (> 5,500 titles)
    - Social Sciences (> 2,800 titles)

Obr. 10 - Úvodní obrazovka databáze Scopus s rozhraním jednoduchého vyhledávání

Úvodní obrazovka nabízí tradičně jednoduché vyhledávání obohacené o možnosti omezení na časovou periodu, typ dokumentu a předmětnou oblast. V dalších záložkách jsou k dispozici specifická vyhledávání, a to autorů, pracovišť a vyhledávání pokročilé. Po provedení vyhledávání jsou k dispozici výsledky podle kategorie zdroje – přímo z databáze Scopus, potom z databáze referencí, které nemají primární záznam v databázi\*, výsledky z webu, které jsou získány pomocí vyhledávače odborné literatury Scirus ([www.scirus.com](http://www.scirus.com)), z patentových databází a z tzv. vybraných zdrojů, tj. z institucionálních repositářů, preprintových archivů, open access archivů aj. Databáze také obsahuje tzv. Articles in Press, tj. články, které jsou před oficiálním publikováním, ale již jsou přijaty do časopisu.

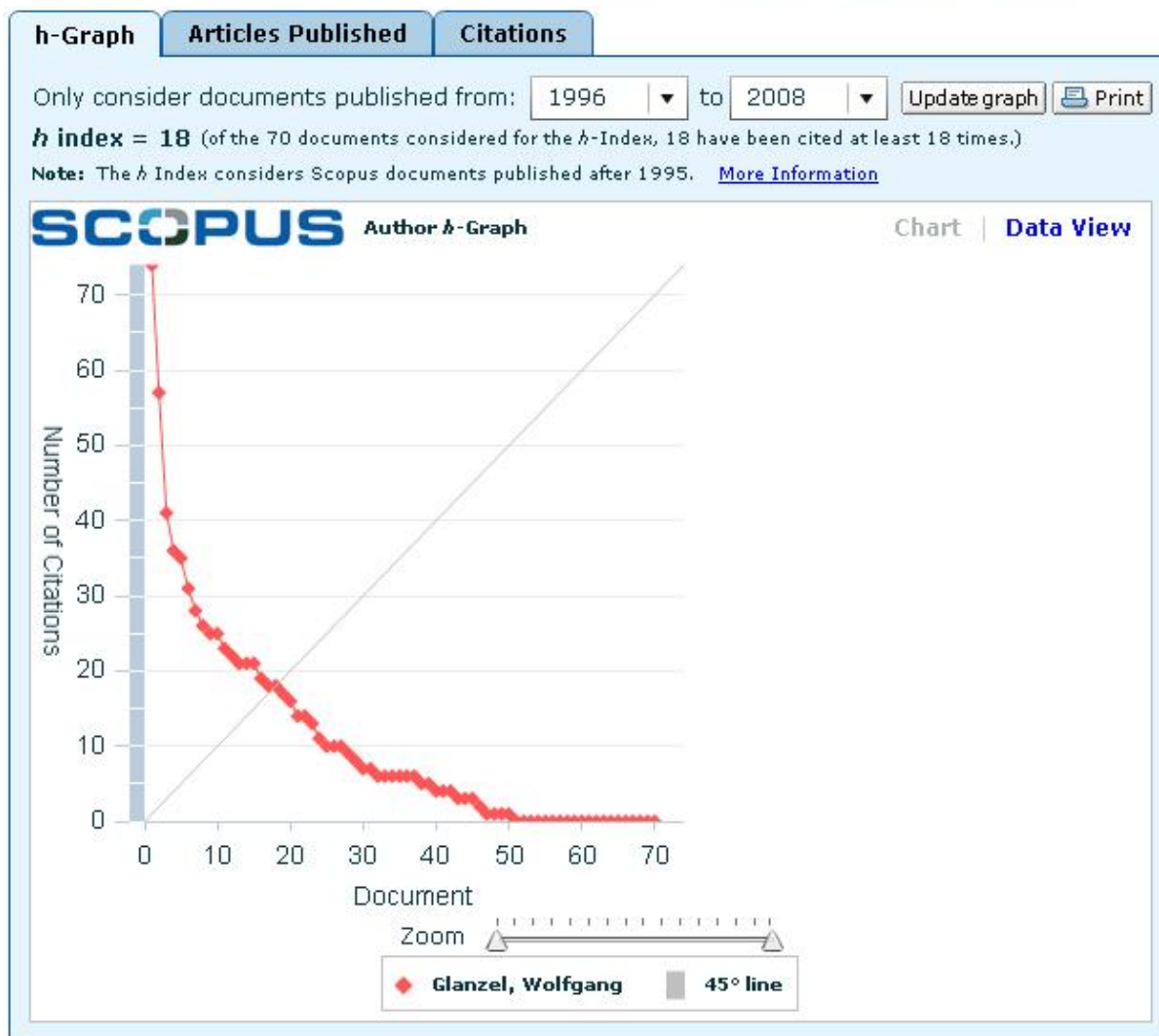
U každého záznamu je uveden počet citací, jakákoliv sada výsledků se dá obdobně jako v ISI Web of Science analyzovat podle let a přijatých citací. U článku je k dispozici seznam referencí přímo opatřený počty citací a dalším aparátem, který umožňuje s tímto seznamem či jednotlivou referencí rovnou pracovat.

Co se týče autorů a pracovišť, v databázi Scopus se nachází jejich autoritní seznamy; pro každého autora či instituci jsou tak zvláště zpracovávány analytické údaje. V profilu autora jsou pak k dispozici poměrně detailní informace, jako variantní názvy jména, afiliace aktuální i historické, publikační historie, odkazy na všechny články, předmětná oblast autora, či seznam spoluautorů. Z analytických dat je k dispozici počet článků, referencí a citací, h-index autora a přehled článků a citací za celou publikační historii podle jednotlivých let. Data autora lze jednoduše analyzovat na tzv. h-grafu, který graficky znázorňuje přehled citací podle článků. K dispozici jsou také grafy publikovaných článků v jednotlivých letech a totéž pro citace. H-graf se dá aplikovat na jakoukoliv sadu výsledků, tedy i na sadu článků, časopis aj.

\* Záznamy, které nejsou původně v databázi, ale objevily se tam právě jako reference, se nazývají „orphan records“.

**Author:** Glanzel, Wolfgang

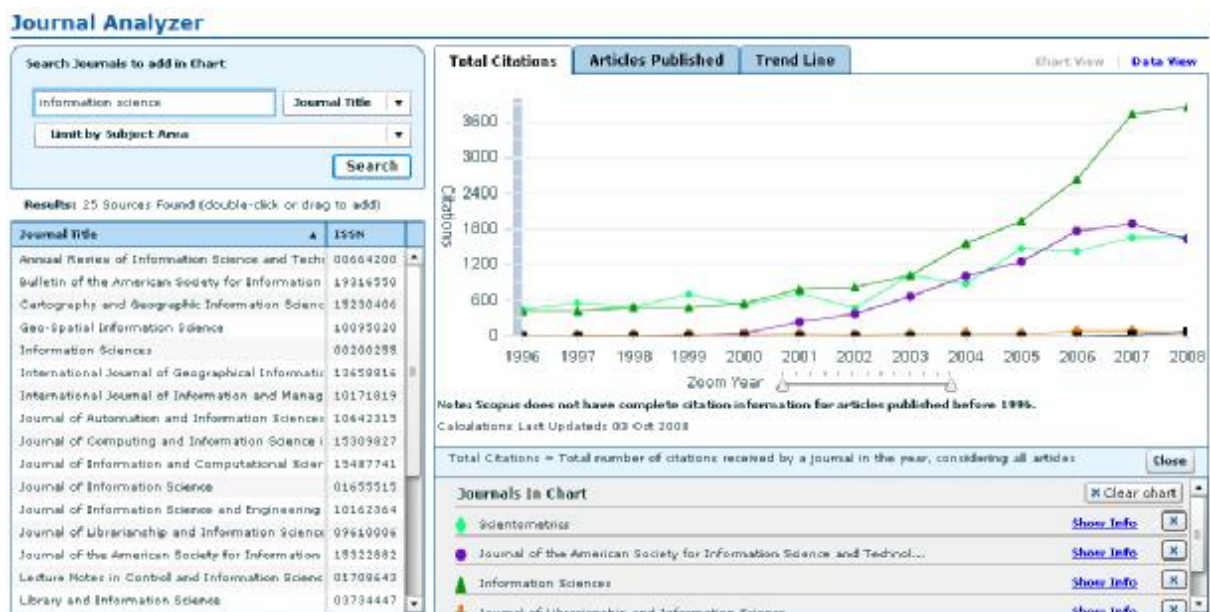
The *h*-Graph measures an author's output and shows the number of citations per document.



**Obr. 11 – Analytický nástroj H-graf, který zobrazuje vztah článků a počtu citací, na průniku křivky grafu s diagonálou lze odečíst H-index**

Vraťme se k pracovištím, jejich identifikaci a dostupným analytickým nástrojům. Mimo popisných informací zde nalezneme údaje jako výčet článků, autorů, počtu patentů, zdrojů, ve kterých autoři tohoto pracoviště publikují, je zpracována analýza spolupracujících institucí. Na koláčovém grafu je zobrazeno rozvržení činnosti v jednotlivých vědeckých disciplínách, které lze přepnout na tabulkové zobrazení.

Pro analytické zpracování časopisů je k dispozici Journal Analyzer v části Analytics, pomocí něhož lze časopisy na základě citačních analýz porovnávat.



Obr. 12 – Analyzátor časopisů, umožňuje vzájemně porovnávat citace a publikované články v jednotlivých časopisech

U profilů jednotlivých časopisů je pak k dispozici přehled všech čísel a k tomu analytický nástroj pro zobrazení citačních analýz dle jednotlivých let.

Databáze nabízí také možnosti personalizace, umožňuje uložit si vlastní seznam autorů, časopisů a článků. K dispozici je možnost zřídit si upozornění na nová čísla časopisů, změnu výsledků vyhledávání a na citace dokumentů nebo autora. Je možné si ukládat historii vyhledávání.

Obdobně jako ISI Web of Science je k dispozici propojení s referenčními manažery (RefWorks), dalšími vyhledávači, katalogy místní knihovny a dalších zdrojů (také Ulrich's Periodical).

Scopus má v sobě zabudováno mnoho nástrojů, které jsou v případě produktů Thomson Reuters rozdrobeny do jednotlivých databází, je velmi přehledný a intuitivní. Omezením je zmíněné zpracování citačních dat pouze po r. 1996, otázkou je, zda-li při současném rychlém vývoji vědy tato retrospektiva není dostatečná. Pravděpodobně v příštím roce by měla být tato databáze k dispozici i v ČR, výsledky z ní jsou již letos zahrnuty do hodnocení vědy v ČR.



### 5.2.3 Google Scholar

The screenshot shows the Google Scholar search interface. The search bar contains the text 'co citation' and the search button is labeled 'Hledat'. To the right of the search bar are three links: 'Rozšířené vyhledávání služby Scholar', 'Předvolby služby Scholar', and 'Nápověda služby Scholar'. Below the search bar, the results are displayed. The first result is titled 'Scholar Všechny články - **Poslední články**' and shows 'Výsledky 1 - 10 z asi 588 000 na dotaz co citation. (0,17 sekund)'. The first result is 'Co-citation in the scientific literature: A new measure of the relationship between two documents - Get@CU via UIG'. Below this title is the author 'H Small - Journal of the American Society for Information Science, 1973 - doi.wiley.com' and a snippet of the text: 'Page 1. Co-citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents A new form of document coupling called co-citation is ...'. There are three links below the snippet: 'Počet citací tohoto článku: 552 - Související články - Hledání na webu - Get This Item at HCT - Všechny verze (počet: 3)'. The second result is 'Visualizing a discipline: An author co-citation analysis of information science, 1972-1995 - rutgers.edu (PDF) - Get@CU via UIG'. Below this title is the author 'HD White, KW McCain - Journal of the American Society for Information Science, 1998 - doi.wiley.com' and a snippet: 'Page 1. Visualizing a Discipline: An Author Co-Citation Analysis of Information Science, 1972-1995 ... oeuvres, or journals. Author co-citation analysis (ACA) ...'. There are four links below the snippet: 'Počet citací tohoto článku: 290 - Související články - Hledání na webu - Get This Item at HCT - Všechny verze (počet: 16)'. The third result is '[CITACE] A Co-Citation Model of a Scientific Specialty: A Longitudinal Study of Collagen Research - Get@CU via UIG'. Below this title is the author 'HG Small - Social Studies of Science, 1977' and a snippet: '... Downloaded from Page 2. 139 A Co-Citation Model of a Scientific Specialty: A Longitudinal Study of Collagen Research Henry G. Small Author ...'. There are four links below the snippet: 'Počet citací tohoto článku: 68 - Související články - Hledání na webu - Get This in Print at HCT - Všechny verze (počet: 2)'. The entire screenshot is enclosed in a light green border.

Obr. 13 – Výsledky vyhledávání v Google Scholar

Google Scholar Beta (<http://scholar.google.cz/> nebo <http://scholar.google.com/>) byl spuštěn na konci roku 2004 a zpočátku zejména odborná veřejnost na něj pohlížela skrze prsty. Google Scholar prohledává databáze vydavatelů odborné literatury, odborných časopisů, odborných repositářů, preprintových a open access archivů (mezi přispěvatele patří Elsevier, American Chemical Society, PubMed, OCLC, Blackwell, the Institute of Physics, the Nature Publishing Group, Wiley Interscience a Springer atd.). Obsahem pokrývá taktéž projekt Google Books. U větších zdrojů se však často vyskytuje problém, že je zaindexována jen jejich malá část (Elsevier, PubMed). Doposud se Google Scholar nepodařilo doplnit zásadní mezery v obsahu o nejlépe hodnocené časopisy [JACSÓ, 2008].

V oblasti citačních analýz poskytuje Google Scholar pouze počet citujících dokumentů s patřičným odkazem na ně. Vyhodnocuje ještě klíčové autory (cca 5) – nejčastěji se vyskytující autory v rámci vyhledaných výsledků. Vzhledem k obrovskému obsahu počty citací nemusí být vypovídající, neboť Google Scholar obsahuje velké množství duplicit, triplicit a vyšších řádů četnosti. To plyne především z nedokonalé práce se strukturovaným textem, stále se objevují neexistující jména autorů, informace v polích, které tam zjevně nepatří, a to často i pokud jsou ve zdrojovém dokumentu uvedena správně. Google Scholar navíc nedisponuje možností seřadit jakkoliv jinak výsledky vyhledávání, jsou seřazeny podle počtu citací [JACSÓ, 2008].

Výhodou Google Scholar je především velmi rychlé jednoduché vyhledávání v obrovském množství zdrojů nehledě na vědeckou disciplínu. Dokáže najít plný text článku, který z klasických databází není dostupný, neboť nemáte zdroj předplacen, na internetu však je. Nabízí také spojení s českými SFX službami a dalšími knihovními katalogy. Velkou výhodou je šíře geografického a jazykového pokrytí. Do Google Scholar se může zapojit takřka kdokoli.

Google Scholar je zdarma, nabízí velmi rychlé vyhledání napříč obory, v současné době ale stále není vhodný pro jakékoliv bibliometrické či scientometrické hodnocení [JACSÓ, 2008].

## 5.3 DALŠÍ ANALYTICKÉ NÁSTROJE A NÁSTROJE MAPOVÁNÍ VĚDY

V dnešní době rozvinutých webových služeb máme k dispozici portály, které nám předkládají automaticky výsledky citačních analýz. Většinou jsou spojené s producentem citační databáze. Mezi takové právě patří portál Scopus TopCited a ISI HighlyCited.com.

### **Scopus TopCited** (<http://www.topcited.com/>)

Portál TopCited nabízí přesně to, co evokuje jeho název. Na základě dat z databáze Scopus zveřejňuje přehledně 20 nejcitovanějších článků pro každou vědeckou disciplínu mezi lety 2004 a 2008 ve zvolené časové periodě. Pomocí tohoto nástroje se velmi snadno získají charakteristiky vědecké disciplíny podle počtu citací. V současné době jsou nejcitovanější články z oblasti genetiky, částicové fyziky a medicíny.

### **ISI HighlyCited** (<http://isihighlycited.com/>)

Tento portál poskytuje profily nejcitovanějších autorů a nejprestižnějších pracovišť, vychází z dat Thomson Reuters a databáze ISI Web of Science. Na tomto portálu se výrazněji projevuje původní proamerické zaměření databáze ISI Web of Science, nenalezneme na něm např. ani jednoho českého vědce.

### **SciMago** (<http://www.scimagojr.com/>)

SCImago Journal & Country Rank je portál, který obsahuje a zpřístupňuje scientometrické indikátory časopisů a zemí vytvořených z dat, která pochází z databáze Scopus. Tato platforma získala název podle indikátoru SCImago Journal Rank (SJR), který byl vyvinut na stejném principu jako je algoritmus Google PageRank [HANUŠOVÁ, 2008]. Tento indikátor poskytuje představu o „viditelnosti“ či „známosti“ vědeckých periodik. Data mají retrospektivu do roku 1996.

U časopisů a států jsou přehledově k dispozici údaje SCImago Journal Rank, H-index, celkové počty dokumentů, referencí, citací, citovatelných dokumentů, porovnání dokumentů a citací, referencí a dokumentů. Časopisy lze filtrovat dle země nebo oboru periodika. Pominout nelze zajímavou stránku World Report, kde jsou sumarizovány údaje o celé světové vědecké produkci. Můžeme se tak například dozvědět, že poměr všech a citovatelných dokumentů je 10:9 či že nejvíce se publikuje v medicíně (cca 20 % dokumentů).

Jednou z hlavních funkcí webu SCImago je možnost porovnání dvou až čtyřech periodik či zemí. Samozřejmostí je možnost omezit porovnání na určitý obor nebo srovnat region (např. pozici České republiky v rámci vědecké výkonnosti ve východní Evropě). K tomu jsou k dispozici i pokročilé vizualizační funkce.

### **Eigenfactor.org** (<http://www.eigenfactor.org/index.php>) [VAVŘÍKOVÁ, 2008]

Portál *Eigenfactor.org* s podtitulem „Ranking and Mapping Scientific Knowledge“ poskytuje hodnocení odborných časopisů, zajímavé vizualizační, statistické a srovnávací výstupy v mapování vědy a velmi srozumitelně ukazuje provázanost vědeckých disciplín. Portál taktéž poskytuje informace o ceně časopisu (v rámci subskripcí) a porovnávání cen časopisů (ve spolupráci s portálem Journalprices.com). Na portálu jsou k dispozici statistiky nejlepších časopisů.

Tento portál analyzuje data poskytnutá Thomson Reuters Journal Citation Reports Dataset, přebírá data z rejstříků Science Citation Index a Social Science Citation Index. K hodnocení užívá pětileté citační okno, tj. hodnocení v r. 2006 je vypočítáno z dat za léta 2001-2005.

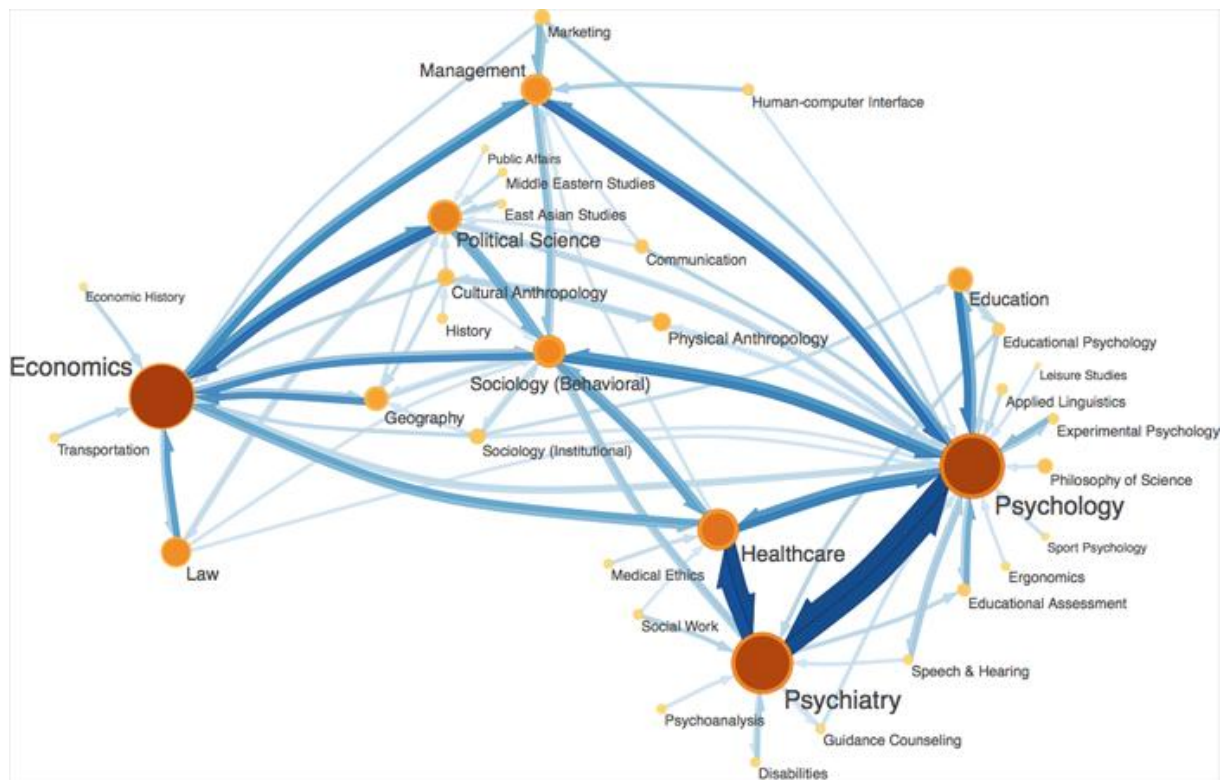
Portál přisuzuje časopisům dva indikátory – Eigenfactor Score a Article Influence. *Eigenfactor Score* (dále ES) je míra důležitosti časopisu pro vědeckou komunitu. Podle definice a matematického modelu je to odhad procenta času, který stráví uživatelé s daným časopisem. Také proto se toto skóre zvětšuje s velikostí časopisu, je přímo závislé na počtu článků, větší časopisy mají také větší ES. Součet všech ES všech časopisů je roven 100, každému časopisu je přisouzena poměrná část ze 100. V r. 2006 měl nejvyšší ES časopis Nature, a to 2.138; v první tisícovce časopisů ES neklesá pod 0.01.

Abychom mohli odhadnout důležitost časopisu nehlédě na jeho velikost (počet publikovaných článků), je zaveden druhý indikátor, *Article Influence* (dále AI). AI je míra průměrného vlivu každého článku v časopise za pět let po jeho publikování. Pokud je AI u časopisu např. 20, znamená to, že průměrný článek v tomto časopise je dvacetkrát vlivnější než průměrný článek v celé kolekci dat. Tento indikátor je srovnatelný s impakt faktorem.

Pro každý časopis jsou k dispozici informace o něm, hodnoty indikátorů portálu a jejich srovnání s impakt faktorem. Grafy názorně zobrazují vývoj časopisu, z čehož v případě porovnání nejlepších časopisů lze usuzovat na vývoj disciplíny jako takové v rámci celé vědy, případně obráceně zjišťovat trendy ve vědě. Statistiky jsou dostupné pro několik let.

Velmi zajímavé jsou vizualizační funkce portálu. Na základě propojení dokumentů citacemi je k dispozici interaktivní mapa vědy, která vždy zobrazuje vazby na příbuzné disciplíny, včetně typu vazby. Druhá vizualizace slouží názornému porovnávání časopisů a analytických údajů.





Obr. 14 – Mapa společenských věd podle portálu Eigenfactor.org

## POUŽITÁ LITERATURA

- BATISTA, P.D., CAMPITELI, M.G., KINOUCI, O. Is it possible to compare researchers with different scientific interests? *Scientometrics*. 2006, vol. 68, no. 1, s. 179–189.
- BJÖRNEBORN, Lennart, INGWERSEN, Peter. Perspectives of webometrics. *Scientometrics*. 2001, vol. 50, no. 1, s. 65-82.
- BJÖRNEBORN, Lennart, INGWERSEN, Peter. Toward a basic framework for webometrics. *Journal of American Society of Information Science and Technology*. 2004, vol. 55, is. 14, s. 1216-1227.
- BOLLEN, Johan, RODRIGUEZ, Marko. A., SOMPEL, Herbert Van de. Journal Status. *Scientometrics*. 2006, vol. 69, no. 3, s. 669-687.
- BORNMANN, Lutz, DANIEL, Hans-Dieter. What Do We Know About the h Index?. *Journal of the American Society for Information Science and Technology*. 2007, vol. 58, is. 9, s. 1381-1385.
- Cybermetrics : international journal of scientometrics, informetrics, and bibliometrics* [online]. Consejo Superior de Investigaciones Científicas, Centro de Informacion y Documentacion Cientifica. 1997- , vol. 1, is. 1- . Spain : 1x ročně. Dostupný z WWW: <<http://www.cindoc.csic.es/cybermetrics/cybermetrics.html>>. ISSN 1137-5019.
- DIODATO, Virgil Pasquale. *Dictionary of Bibliometrics*. New York : Haworth Press, 1994. 185 s. Dostupný z WWW: <<http://books.google.cz/books?id=XBg1SNzNTD0C&printsec=frontcover>>. ISBN 1560248521.
- EGGHE, Leo. From h to g: the evolution of citation indices. *Research Trends* [online]. 2007, is. 1 [cit. 2008-08-24]. Dostupný z WWW: <[http://info.scopus.com/researchtrends/archive/RT1/exp\\_op\\_1.htm](http://info.scopus.com/researchtrends/archive/RT1/exp_op_1.htm)>.
- GLÄNZEL, Wolfgang. On the h-index : A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*. 2006, vol. 67, no. 2, s. 315-321.
- GLÄNZEL, Wolfgang, DEBACKERE, Koenradd. *Measuring communication in science : Opportunities and limitations of bibliometric methods* [online]. Leuven : Steunpunt O&O Indicatoren, [2004] [cit. 2008-09-06]. Ppt

prezentace. Dostupný z WWW: [http://www.google.cz/url?sa=t&source=web&ct=res&cd=1&url=http%3A%2F%2Fwww.ulb.ac.be%2Funica%2Fdocs%2FSch-com-2004-pres-Glanzel.ppt&ei=9qfDSPCBJSa1wbM74ToBw&usg=AFQjCNH15XCttR2zAwIF-Q2ZCXcS5vz1EQ&sig2=d\\_awnWCYJSv0eeZ\\_PI3bng](http://www.google.cz/url?sa=t&source=web&ct=res&cd=1&url=http%3A%2F%2Fwww.ulb.ac.be%2Funica%2Fdocs%2FSch-com-2004-pres-Glanzel.ppt&ei=9qfDSPCBJSa1wbM74ToBw&usg=AFQjCNH15XCttR2zAwIF-Q2ZCXcS5vz1EQ&sig2=d_awnWCYJSv0eeZ_PI3bng).

11. GODIN, B. On the origins of bibliometrics. *Scientometrics*. 2001, vol. 68, no. 1, s. 109-133.
12. GRANOVSKY, Y. V. Is it possible to measure science? V. V. Nalimov's research in scientometrics. *Scientometrics*. 2001, vol. 52, no. 2, s. 127-150.
13. HANUŠOVÁ, Soňa. SCImago Journal & Country Rank. *Ikaros* [online]. 2008, roč. 12, č. 6 [cit. 2008-10-17]. Dostupný z WWW: <http://www.ikaros.cz/node/4829>. URN-NBN:cz-ik4829. ISSN 1212-5075.
14. HJØRLAND, Birger. *Impact Factor* [online]. 2007, 21-4-2007 [cit. 2008-08-24]. Dostupný z WWW: [http://www.db.dk/bh/core%20concepts%20in%20lis/articles%20a-z/impact\\_factor.htm](http://www.db.dk/bh/core%20concepts%20in%20lis/articles%20a-z/impact_factor.htm).
15. INGWERSEN, Peter, LARSEN, Birger. *Advanced publication and citation analysis* [ppt prezentace]. Danmarks Biblioteksskole. 2007, [cit. 2008-09-20]. Studijní materiál ke kurzu Information Seeking and Information Retrieval Theories.
16. JACSÓ, Péter. Google Scholar Revisited. *Online Information Review* [online]. 2008, vol. 32, is. 1 [cit. 2008-10-10]. ISSN 1468-4527.
17. JANOŠ, Karel. 2001. Informační etika. In PAPIK, Richard; SOUČEK, Martin; STÖCKLOVÁ, Anna. *Informační studia a knihovnictví v elektronických textech I* [CD-ROM]. Praha : Ústav informačních studií a knihovnictví FF UK, c2001.
18. JANSSENS, Frizo, et al. Towards mapping library and information science. *Information Processing and Management*. 2006, vol. 42, is. 6, s. 1614-1642.
19. KÖNIGOVÁ, Marie. 2001. Vybrané kapitoly z informační vědy. In PAPIK, Richard; SOUČEK, Martin; STÖCKLOVÁ, Anna. *Informační studia a knihovnictví v elektronických textech I* [CD-ROM]. Praha : Ústav informačních studií a knihovnictví FF UK, c2001.
20. KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna České republiky, 2003. Dostupný z WWW: <http://sigma.nkp.cz/cze/ktd>.
21. MOED, Henk F. *Citation Analysis in Research Evaluation*. [s.l.] : Springer, c2005. 346 s. Information Science and Knowledge Management; sv. 9. ISBN 1-4020-3713-9.
22. NORUZI, A. The web impact factor : A critical review. *Electronic Library*. 2006, vol. 24, is. 4, s. 490-500.
23. OKUBO, Y. Bibliometric indicators and analysis of research systems: Methods and examples. OECD, Paris, 1997, STI Working Papers 1997/1 – OCDE/GD(97)41.
24. SEGLEN, Per O. Citations and Journal Impact Factors : Questionable Indicators of Research Quality. *Allergy : European Journal of Allergy and Clinical Immunology*. 1997, Vol. 52, no. 11, s. 1050-1056.
25. SHAPIRO, F.R. Origins of bibliometrics, citation indexing and citation analysis: the neglected legal literature. *Journal of the American Society for Information Science and Technology*. 1999, vol. 43, is. 5, s. 337-339.
26. SCHWARZ, Josef. Před 40 lety začala zlatá éra citačních rejstříků vědecké literatury. *Ikaros* [online]. 2003, roč. 7, č. 11 [cit. 2008-02-23]. Dostupný z WWW: <http://www.ikaros.cz/node/1520>. URN-NBN:cz-ik1520. ISSN 1212-5075.
27. Scientometrics from past to present. *Research Trends* [online]. 2007, is. 1 [cit. 2008-09-06]. Dostupný z WWW: [http://info.scopus.com/researchtrends/archive/RT1/bib\\_mes\\_1.htm](http://info.scopus.com/researchtrends/archive/RT1/bib_mes_1.htm).
28. *SCImago: Journal & Country Rank* [online]. c2007-2008 [cit. 2008-10-10]. Dostupný z WWW: <http://www.scimagojr.com/>
29. *Scopus* [online]. c2008 [cit. 2008-10-10]. Dostupný z WWW: <http://www.scopus.com/>.
30. SMELSER, N. J., BALTES, P. B. *International Encyclopedia of the Social & Behavioral Sciences*. Loet Leydesdorff. Amsterdam : Elsevier, c2004. Scientometrics, s. 13752-13755. ISBN 0-08-043076-7.

31. SOUČEK, Martin. *Informační věda : studijní materiály k předmětu \ "Úvod do informační vědy\ "* [online]. 2006 , 17.10.2006 [cit. 2008-09-06]. Dostupný z WWW: <<http://uisk.ff.cuni.cz/detail.do?articleId=1959>>.
32. ŠPÁLA, Milan. Impakt faktor – Dobrý sluha, ale špatný pán. *Ikaros* [online]. 2006, roč. 10, č. 4 [cit. 2008-08-24]. Dostupný z WWW: <<http://www.ikaros.cz/node/3293>>. URN-NBN:cz-ik3293. ISSN 1212-5075.
33. TAGUE-SUTCLIFFE, Jean. An introduction to informetrics. *Information Processing & Management*. 1992, vol. 28, is. 1, s. 1-3.
34. *The effect of open access and downloads ('hits') on citation impact: a bibliography of studies* [online]. 4 August 2008 [cit. 2008-08-23]. Dostupný z WWW: <<http://opcit.eprints.org/oacitation-biblio.html>>.
35. *Thomson Reuters* [online]. c2008 [cit. 2008-10-10]. Dostupný z WWW: <<http://www.thomsonreuters.com/>>.
36. VAVŘÍKOVÁ, Lucie. Hodnocení a mapování vědy pomocí portálu Eigenfactor.org. *Ikaros* [online]. 2008, roč. 12, č. 8 [cit. 2008-10-17]. Dostupný z WWW: <<http://www.ikaros.cz/node/4935>>. URN-NBN:cz-ik4935. ISSN 1212-5075.
37. VINKLER, Peter. An attempt for defining some basic categories of scientometrics and classifying the indicators of evaluative scientometrics. *Scientometrics*. 2001, vol. 50, no. 3, s. 539-544.
38. XING, Wenpu, GHORBANI, Ali. Weighted PageRank Algorithm. In *Second annual conference on communication networks and services research CNSR'04* . Fredericton : [s.n.], 2004. s. 305-314. Dostupný z WWW: <[http://ias.cs.unb.ca/papers/xingw\\_weighted.pdf](http://ias.cs.unb.ca/papers/xingw_weighted.pdf)>.