

**UNIVERZITA SV. CYRILA A METODA V TRNAVE**  
**FAKULTA PRÍRODNÝCH VIED**  
**KATEDRA BIOLÓGIE**

# **Diplomová práca**

**TRNAVA 2016**

**JURAJ SZÁSZ**



**UNIVERZITA SV. CYRILA A METODA V TRNAVE**  
**FAKULTA PRÍRODNÝCH VIED**  
**KATEDRA BIOLÓGIE**

---

# **Scientometrická analýza FPV UCM v Trnave**

Diplomová práca

**Juraj Szász**

**Školiteľ: prof. RNDr. Ján Kraic, PhD**

**Trnava 2016**

# Bibliografický záznam

<b>Autor:</b>	Bc. Juraj Szász Fakulta prírodných vied, Univerzita sv. Cyrila a Metoda v Trnave Katedra biológie
<b>Názov práce:</b>	Scientometrická analýza Fakulty prírodných vied Univerzity sv. Cyrila a Metoda v Trnave
<b>Študijný program:</b>	Aplikovaná biológia
<b>Študijný obor:</b>	Biológia
<b>Školiteľ</b>	prof. RNDr. Ján Kraic, PhD
<b>Akademický rok:</b>	2016/2017
<b>Počet strán:</b>	xiii + 40
<b>Kľúčové slová:</b>	Kľúčové slovo; Kľúčové slovo; Kľúčové slovo; Kľúčové slovo; Kľúčové slovo; Kľúčové slovo; Kľúčové slovo; Kľúčové slovo

# Bibliographic Entry

<b>Author:</b>	Bc. Juraj Szász Faculty of Natural Science, University of Ss. Cyril and Methodius in Trnava Department of Biology
<b>Title of Thesis:</b>	Scientometric analysis of the Faculty of Natural Sciences, University of Ss. Cyril and Methodius in Trnava
<b>Degree Programme:</b>	Applied Biology
<b>Field of Study:</b>	Biology
<b>Supervisor:</b>	prof. RNDr. Ján Kraic, PhD
<b>Academic Year:</b>	2016/2017
<b>Number of Pages:</b>	xiii + 40
<b>Keywords:</b>	Keyword; Keyword; Keyword; Keyword; Keyword; Keyword; Keyword; Keyword; Keyword

# Abstrakt

V tejto diplomovej práci sa venujeme štúdiu Neviditeľného ružového jednorožca. Kľúčovou vlastnosťou tejto ružovej bytosti je to, že je *ad hoc* neviditeľná a jej neexistenciu nie je možné dokázať. Vieme, že je ružová, pretože to tvrdia ľudia, ktorým sa zjavila. V našej práci sa nám podarilo dokázať, že Neviditeľný ružový jednorožec stvoril vesmír po smrti bude každého človeka súdiť podľa toho, čo počas života robil v spálni. Tento záver nie je možné vyvrátiť, čo znamená že to musí byť absolútna pravda.

# Abstract

In this thesis we study Invisible Pink Unicorn. Key property of this pink being is that He is *ad hoc* invisible and His nonexistence is unprovable. We know that He is pink, because it was said by the people to whom He has revealed. We have managed to prove that the Invisible Pink Unicorn is the Creator of the universe and every human will be judged by Him after death according to what he/she were doing in bedroom. This conclusion is not falsifiable and it means that it has to be the absolute true.

# Pod'akovanie

Na tomto mieste by som sa chcel poďakovať in memoriam Richardovi P. Feynmanovi za to, že bol pre mňa celý čas bezodným zdrojom inšpirácie. Ďalej by som sa chcel poďakovať Gabrielovi Szászovi za vytvorenie šablóny pre typografický systém L<sup>A</sup>T<sub>E</sub>X v ktorom bola napísaná táto práca. V neposlednom rade by som sa chcel poďakovať výrobcovi tabletiiek AlphaD3 bez ktorých by táto práca nikdy nevznikla.

# Prehlásenie

Prehlasujem, že som svoju diplomovú prácu vypracoval samostatne s využitím informačných zdrojov, ktoré sú v práci citované.

Trnava, 1. decembra 2016

.....  
Juraj Szász

# Obsah

<b>Úvod</b> .....	<b>viii</b>
<b>Prehľad použitých symbolov</b> .....	<b>ix</b>
<b>Kapitola 1. Literárny prehľad</b> .....	<b>1</b>
1.1 Definícia pojmov .....	1
1.1.1 Bibliometria .....	1
1.1.2 Scientometria .....	2
1.1.3 Informetria .....	2
1.1.4 Webometria .....	2
1.1.5 Cybermetria .....	2
1.1.6 Bibliometrické zákony .....	2
1.1.7 Lotkov zákon .....	3
1.1.8 Bradfordov zákon .....	3
1.1.9 Zipfov zákon .....	4
1.2 Citačné registre .....	4
1.2.1 Web of Science (WoS) .....	5
1.2.2 Scopus .....	7
1.3 Citačné indikátory .....	8
1.3.1 Journal Impact Factor (IF) .....	8
1.3.2 Hirshov index ( $h$ -index) .....	9
1.3.3 Eggheov index ( $g$ -index) .....	9
1.3.4 Zhangov $e$ -index .....	9
1.3.5 Súčasný $h$ -index (Contemporary $h$ -index) .....	10
1.3.6 Citačná frekvencia váhovaná podľa veku (Age-weighted citation rate – AWCR) a AW-index .....	10
1.3.7 Individuálny (Individual) $h$ -index .....	11
1.3.8 Individuálny (Individual) $h$ -index $h_{I,norm}$ .....	11
1.3.9 Multi-autorský $h$ -index .....	11
<b>Kapitola 2. Druhá kapitola s matematikou <math>\int f(x) dx</math> v názvu</b> .....	<b>12</b>
2.1 Podkapitola .....	12
<b>Kapitola 3. Druhá kapitola s matematikou <math>\int f(x) dx</math> v názvu</b> .....	<b>13</b>
3.1 Podkapitola .....	13

<b>Záver</b> .....	<b>14</b>
<b>Príloha</b> .....	<b>15</b>



# Úvod

Scientometria (ang. *scientometrics*) je vedný obor, ktorý sa zaoberá hodnotením vedy, t.j. vedeckých publikácií, vedeckých pracovníkov a vedeckého pokroku použitím matematických, štatistických metód. Súbor týchto metód sa nazýva *scientometrika*. Hlavným aspektom, podľa ktorého sa hodnotia vedecké práce sú citácie, t.j. referencie na iné publikácie, ktoré autor použil, alebo chce na ne upozorniť. Všeobecne je brané, že publikácia, ktorú cituje viac iných vedeckých článkov má väčší impakt (dopad). To znamená, že práca je populárna, používaná, pretože je kvalitná a prínosná pre vedecký pokrok.

Ďalším spôsobom hodnotenia vedy je použitie ekonomických aspektov. Počet a hodnota grantov, ktoré daný pracovník, či inštitúcia dosiahli, alebo hodnota praktického uplatnenia konkrétnych poznatkov. Bohužiaľ tento spôsob hodnotenia nemôže byť všeobecný a ani spravodlivý, pretože cieľom vedy niekedy nemusí primárne vytvoriť prospešný a ekonomicky výhodný produkt, ale posunúť ľudské poznanie. Väčšina výskumu patrí do tzv. základného výskumu, v ktorom sa bezprostredne neočakáva možnosť aplikácie nadobudnutých poznatkov do praxe. Mnohé z nich nie sú doteraz aplikovateľné a niektoré našli praktické uplatnenie až po uplynutí niekoľko storočí (napr. matematické modely umelej inteligencie). V neposlednom rade predmetom výskumu je overenie hypotézy. Vedec by mal očakávať, že výsledok výskumu bude vyvrátenie hypotézy a automaticky neprinesie ekonomický úžitok, ale iba pokrok. Každý vedec musí mať na pamäti, že aj negatívny výsledok je výsledok.

Cieľom tejto práce je scientometrické hodnotenie publikačnej činnosti zamestnancov Fakulty prírodných vied Univerzity sv. Cyrila a Metoda v Trnave. Hodnotenie je vykonané kvantitatívne, počtom publikovaných prác a tiež kvalitatívne pomocou tzv. citačných indexov vypočítaných programom *Publish or Perish*. Vstupné dáta do hodnotenia boli získané z najväčších a najvýznamnejších citačných databáz *Elsevier Scopus* a *Thomson Reuters Web of Science*.

# Prehľad použitých symbolov

$IF$	Journal Impact Factor
$h$	Hirshov index ( $h$ -index)
$AR$	citačná frekvencia váhovaná podľa veku
$f(n)$	počet autorov, ktorí publikovali $n$ článkov v danom obore
$h_i$	individuálny Hirshov index ( $h_i$ -index)
$h_{I,norm}$	normovaný $h_i$ -index (program <i>Publish or Perish</i> )
$h_m$	multiautorský Hirshov index ( $h_m$ -index)
$r_{eff}(r)$	efektívny rang článku $r$

# Kapitola 1

## Literárny prehľad

### 1.1 Definícia pojmov

#### 1.1.1 Bibliometria

Termín bibliometria je zložený z dvoch gréckych slov: biblion, čo zamená kniha a métron, meranie. Takže doslovný preklad by bol meranie kníh, alebo veda zaoberajúca sa meraním kníh. Zrozumiteľnejšia je prvá definícia: aplikácia matematických a štatistických metód na knihy a iné komunikačné média. [Pritchard \(1969\)](#)

V súčasnosti sa pod týmto termínom chápe súhrn štatistickým metód na kvantitatívnu analýzu publikácií v písomnej forme, ako sú knihy, alebo články vo forme bibliografických záznamov. Tieto záznamy zahrňujú informácie ako názov publikácie, jej autorov, rok publikovania, ale aj kľúčové slová, abstrakt, či referencie na iné publikácie. na bibliometrických záznamoch môžeme študovať:

- aspekty tvorby publikácií – autori, použitá literatúra,
- aspekty šírenia publikácií – komunikačné kanály ako názov časopisu,
- aspekty použitia publikácií – citačné prepojenia, ale aj štatistika požičiavania v knižnici, alebo frekvencia prístupu cez web. ([Ondrišová, 2011](#))

Najčastejšia bibliometrická metóda je tzv. citačná analýza, pri ktorej sa štatisticky spracovávajú citačné prepojenia na iné dokumenty (citácie). V nej sa ďalej zahrňujú ostatné informácie bibliografických záznamov, ako počet autorov (priemerný počet autorov na dokument, priemerný počet citácií na autora za dokument), počet strán (priemerný počet citácií na stranu dokumentu), počet publikácií v konkrétnom časopise a zmeny týchto informácií za isté obdobie. To znamená, že analýzou dát z bibliometrických záznamov môžeme sledovať vývoj jednotlivých oblastí, ich vzájomný vplyv a prepojenia.

Na základe týchto empirických dát sa vytvárajú matematicko-štatistické modely, ktorými sa snažia opísať procesy súvisiace s tvorbou, šírením a použitím zaznamenaných informácií.

Bibliometria úzko súvisí s ďalšími disciplínami ako scientometria, informetria, librametria, webometria a cybermetria. Všetky tieto disciplíny skúmajú kvantitatívne aspekty informácií a preto je metodika veľmi podobná, líšia sa iba oblasťou, ktorú skúmajú.

### 1.1.2 Scientometria

Termín scientometria môžeme rozdeliť na dve slová: latinské *scientia*, čo znamená poznanie a už spomínané grécke *métron*, teda meranie. doslovne „meranie poznania.“ Pojem scientometria Nalimov (1969) definoval ako kvantitatívne metódy, ktoré sú používané na analýzu vedeckého poznania a výskumu.

Scientometriu je možné považovať ako aplikáciu bibliometrie na vedecký výskum a pokrok. V súčasnosti sa na kvantifikáciu vedeckého pokroku využívajú vedecké články. Lenže ich samotný počet nič nehovorí o ich kvalite. Indikátorom kvality vedeckých publikácií sú tzv. citácie. Teda odkazy na pôvodnú publikáciu, z ktorej čerpajú. Ich počet je kvantitatívnym znakom kvality článku. Pri analýze niekoľkých článkov, napr. vyprodukovaných jedným pracovníkom je potrebné zahrňovať distribúciu citácií, medzi článkami. Na to slúži tzv. citačná analýza.

Okrem vedeckých publikácií scientometria skúma aj ďalšie kvantitatívne aspekty vedy ako napr. človekoroky, počet rokov praxe vedcov, finančné vstupy apod. [Bellis \(2009\)](#)

### 1.1.3 Informetria

Pod termínom informetria sa chápu kvantitatívne aspekty informácií v ľubovoľnej forme v ľubovoľnej sociálnej skupine.

Termín sa začal používať až koncom 80-tych rokov ako spoločný názov pre bibliometriu a scientometriu, ale stále sa bibliometria, scientometria a informetria používajú ako synonymá.

### 1.1.4 Webometria

S rozvojom informačných technológií a hlavne internetu sa presunula pozornosť na informácie v prostredí internetu. Webometria skúma kvantitatívne aspekty konštrukcie a využívania informačných zdrojov, štruktúr a technológií na webe čerpajúc z bibliometrických a informetrických prístupov.

### 1.1.5 Cybermetria

Cybermetria sa na rozdiel od webometrie sa zaoberá kvantitatívnymi aspektami iných internetových služieb ako sú diskusné skupiny, alebo elektronická pošta.

### 1.1.6 Bibliometrické zákony

Pod termínom bibliometrický zákon (alebo taktiež nazývaný informetrický zákon) chápeme matematický model, ktorý opisuje empirické závislosti bibliometrických dát a javy ako distribúciu dokumentov v istom súbore rôznych autorov, alebo distribúciu citácií v istom súbore dokumentov apod. Bibliometrické zákony sú odvodené ako generalizácia istých štatistických dát. (Todeschini et al., 2016)

V období medzi rokmi 1920 a 1930 boli publikované tri hlavné bibliometrické zákony: Lotkov zákon distribúcie vedeckých prác medzi autormi, Bradfordov zákon rozdelenia

publikácií konkrétneho oboru vo vedeckých časopisoch a Zipfov zákon distribúcie slov v texte (De Bellis, 2009).

### 1.1.7 Lotkov zákon

Pomenovaný podľa amerického chemika, matematika a štatistika Alfreda J. Lotku opisuje frekvenciu publikácie prác v danom obore vzhľadom na autorov. Lotka zoradil autorov podľa počtu publikácií a analyzoval koľko prác prislúcha k prvému autorovi, druhému atď. Dáta čerpal z indexov *Chemical Abstract* a *Geschichtstafeln der Physik* (Lotka, 1926). Vyšla mu jednoduchá matematická závislosť. Počet autorov  $f(n)$ , ktorí publikovali  $n$  článkov v danom obore ( $n = 1, 2, 3, \dots$ ) sa blíži ku  $1/n^2$  násobku počtu autorov, ktorí publikovali jeden článok.

Lotkov zákon je matematicky definovaný vzťahom (1.1), v ktorom  $K$  a  $\alpha$  sú kladné konštanty závisiace na vedeckej oblasti. Vo väčšine prípadov platí, že  $\alpha = 2$  a  $K = 1$  (Egghe, 2005).

$$f(n) = \frac{K}{n^\alpha} \quad (1.1)$$

Ak je známy počet autorov s jedným článkom ( $a_1$ ), je možné pomocou vzťahu (1.2) z Lotkovho zákona určiť približný počet autorov s  $n$  publikáciami v danom vedeckom obore.

$$a_n = \frac{a_1}{n^2} \quad (1.2)$$

Napríklad v súbore 100 autorov by 4 autori mali mať každý 5 publikácií ( $100/5^2 = 4$ ).

### 1.1.8 Bradfordov zákon

Britský knihovník Samuel Clement Bradford, si všimol istú pravidelnosť v distribúcii počtu článkov s konkrétnou tematikou vo vedeckých časopisoch. V roku 1934 publikoval prácu, v ktorej popísal tento jav. V danej vedeckej práci študoval bibliografické záznamy časopisov z oblasti geofyziky. Články týkajúce sa istej témy našiel v 326 časopisoch. Potom zostupne usporiadal časopisy podľa počtu článkov spadajúcich do danej témy. Nakoniec ich rozčlenil do troch skupín tak, aby každá skupina obsahovala zhruba taký istý počet článkov. Vyšlo mu:

- prvá skupina obsahovala 9 časopisov s 429 článkami,
- druhá skupina obsahovala 59 časopisov s 499 článkami,
- tretia skupina obsahovala 258 časopisov s 404 článkami.

Prvú skupinu s najväčším počtom článkov na časopis pomenoval ako jadro, druhú pomenoval ako prvú zónu a tretiu pomenoval ako druhú zónu.

Počty časopisov v jednotlivých skupinách dal do pomeru:

$$9 : 59 : 258 \quad (1.3)$$

ktorý sa blíži ku:

$$9 : (9 \cdot 5) : (9 \cdot 5^2) \quad (1.4)$$

Teda pomer:

$$9 : 5 : 5^2 \quad (1.5)$$

Podľa, ktorého definoval všeobecnú matematickú definíciu ako:

$$1 : n : n^2 : \dots \quad (1.6)$$

pričom  $n$  sa nazýva Bradfordov násobok a je závislý od konkrétnych bibliometrických dát.

Bradfordov zákon je považovaný za najlepší model vedeckého výskumu knižničnej a informačnej vedy (Nicolaisen, 2007).

### 1.1.9 Zipfov zákon

Americký jazykovedec George Kingsley Zipf študoval kvantitatívnu analýzu jazyka. Konkrétne analyzoval text knihy *Odyseus* od Jamesa Joycesa. Vybral z textu 29 899 špecifických slov (vylúčil bežné slová ako predložky, spojky apod.) a zoradil ich podľa frekvencie výskytu. Prvé najvfrekventovanejšie slovo dostalo rang 1, druhé rang 2, atď. Potom vynásobil frekvenciu výskytu každého slova s príslušným rangom. Prekvapujúco mu vyšli veľmi podobné hodnoty. Toto zistenie definoval matematicky ako:

$$c = r \cdot f \quad (1.7)$$

pričom  $r$  je rang (poradové číslo) daného slova a  $f$  je frekvencia výskytu slova v texte. Tým pádom  $c$  je konštanta, ktorá reprezentuje daný text (Powers, 1998).

Paradoxne Zipfov zákon neplatí iba v lingvistike, ale je ho možné aplikovať v každej oblasti, kde sa skúma frekvencia výskytu konkrétneho javu. Ako napr. distribúcia počtu citácií, alebo návštevnosť webových stránok (Li, 2002).

Zipfov zákon je možné aplikovať na počty obyvateľov v mestách. V najväčšom meste je dvojnásobok počtu obyvateľov ako v druhom najväčšom meste a trojnásobok ako v treťom najväčšom meste (Jiang et. al, 2015).

## 1.2 Citačné registre

Citačné registre (indexy) sú databázy, z ktorých je možné dohľadať citačné odkazy na publikované odborné texty. Ich analýzou je možné objektívne posúdiť kvalitu citovaných publikácií. Citačné registre vznikli preto, aby bolo možné sledovať, aké ohlasy vo vedeckej komunite vzbudila daná publikácia.

Z počiatku citačné indexy vychádzali v tlačenej forme a ukladané boli ako mikrofilmy. S príchodom nových elektronických médií sa k nim pridali magnetické pásky a CD-ROM nosiče. Od rozšírenia internetu sú všetky citačné registre prístupné on-line.

### 1.2.1 Web of Science (WoS)

*Web of Science* je online platená služba umožňujúca prístup ku citačným registrom a ich citačnú analýzu. Poskytuje komplexné vyhľadávanie vo viacerých citačných a abstraktových databázach, ktoré umožňuje dôkladne scientometricky študovať medziodborové oblasti výskumu. V minulosti mala názov *Web of Knowledge* a bola spravovaná Inštitútom pre vedecké informácie *Institute for Scientific Information* (ISI). V súčasnosti je vo vlastníctve mediálneho gigantu *Thomson Reuters* so sídlom New Yorku, USA (Drake, 2005).

WoS pozostáva z citačných registrov:

- **Science Citation Index Expanded® (SCI-E):**  
Zahrňuje publikácie z viac než 8 500 hlavných časopisov, ktoré pokrývajú 150 vedeckých disciplín od roku 1900 do súčasnosti.
- **Social Sciences Citation Index® (SSCI):**  
Články z viac než 3 000 časopisov 55 oborov sociálnych vied a vybrané publikácie z 3 500 vo svete najdôležitejších vedeckých a technických časopisov od roku 1900 po súčasnosť.
- **Arts & Humanities Citation Index® (A&HCI):**  
Indexuje viac než 1 700 časopisov z oblasti umenia a humanitných vied a vybrané články z viac než 250 časopisov z oblasti sociálnych vied od roku 1975 do súčasnosti.
- **Index Chemicus® (IC):**  
Obsahuje viac než 2,6 milióna záznamov zlúčenín od roku 1993.
- **Current Chemical Reactions® (CCR):**  
Zahrňuje viac než milión chemických reakcií od roku 1986, plus záznamy z francúzskeho Inštitútu duševného vlastníctva (INPI) v časovom rozmedzí od roku 1840 do 1985.
- **Book Citation Index® – Science (BKCI-S) a Book Citation Index® – Social Sciences & Humanities (BKCI-SSH):**  
Poskytuje viac než 50 000 vybraných kníh. 10 000 nových kníh pridávaných každý rok od 2005 po súčasnosť.
- **Conference Proceedings Citation Index® – Science (CPCI-S) a Conference Proceedings Citation Index® – Social Sciences & Humanities (CPCI-SSH):**  
Zahrňuje príspevky z viac než 160 000 konferencií z 256 rôznych oblastí „vedy a techniky (CPCI-S)“ a „sociálnych a humanitných vied (CPCI-SSH)“ od roku 1990. Každým rokom do nej pribúda takmer 400 000 konferenčných príspevkov z cca. 12 000 konferencií.<sup>1</sup>

Zakladateľom prvého moderného citačného registru sa stal Eugen Garfield (1955). V roku 1960 založil inštitúciu *Institute for Scientific Information* (ISI), ktorá od nasledujúceho roku začala vydávať prvý multidisciplinárny „Citačný index pre prírodné vedy“ *Science Citation Index* (SCI). Od roku 1972 sa k nemu pridala „Citačný index pre sociálne

<sup>1</sup>[http://wokinfo.com/products\\_tools/multidisciplinary/webofscience/](http://wokinfo.com/products_tools/multidisciplinary/webofscience/)

vedy“ *Social Science Citation Index* (SSCI) a od roku 1978 „Citačný index pre umenie a humanitné vedy“ *Arts & Humanities Citation Index* (AHCI), (Smith, 2012).

*Web of Science* je časť multidisciplinárnej citačnej databázy *ISI Web of Knowledge*, ktorá je vo vlastníctve spoločnosti *Thomson Reuters*. Mimo už spomínaného citačného registru *WoS* obsahuje:

- **Current Contents Connect** – obsah a bibliografické informácie z viac ako 8 000 vedeckých časopisov,
- **Journal Citation Reports** – ročné bibliometrické hodnotenie a porovnávanie vedeckých časopisov,
- **Essential Scientific Indicator** – hodnotenie a porovnávanie inštitúcií, krajín a vedných oblastí,
- **InCites** – bibliometrické analýzy a hodnotenia inštitúcií, krajín, vedných oblastí,
- **Converis** – komplexný informačný systém výskumnej činnosti univerzít,
- **ScholarOne** – manažérsky systém na peer review časopisov, konferencií a kníh,
- **EndNote** – komplexný nástroj na bibliometrické analýzy.<sup>2</sup>

V rozšírenom vyhľadávaní je používateľovi umožnené detailnejšie formulovať vyhľadávaciu požiadavku pomocou logických operátorov. Výsledný zoznam publikácií môže byť usporiadaný podľa roku vydania, relevancie k danej téme, počtu citácií, názvu časopisu alebo konferencie. Výsledky vyhľadávania je možné dodatočne zúžiť pomocou selekcie konkrétnych predmetových oblastí, typu dokumentu, autora, roku publikovania a pod. Ku každému záznamu je zobrazený aj počet citácií, ktoré publikácia získala, zoznam dokumentov, ktoré ju citovali alebo zoznam podobných dokumentov na základe bibliografického združovania. Ak má používateľ zaplatený prístup, tak je dostupný aj plný text publikácie. Vyhľadané záznamy je možné exportovať v rôznych formátoch. Po exporte sa dajú ďalej analyzovať pomocou bibliometrického softvéru.<sup>3</sup>

Dostupné analytické nástroje umožňujú štatisticky vyhodnocovať vyhľadané záznamy. Výsledky sú prezentované graficky.

Funkcia **Analyze Results** ponúka štatistické vyhodnotenie resp. publikačnú analýzu podľa autorov, roku vydania, predmetovej oblasti, krajiny, jazyka apod.

Funkcia **Create Citation Report** analyzuje publikácie z hľadiska počtu získaných citácií. Zobrazuje počty publikácií v jednotlivých rokoch a rovnako počty citácií, ktoré vybraná množina záznamov získala. Ak by sme hľadali publikácie konkrétneho autora, tak tieto grafy zobrazujú jeho publikačnú produktivitu a úspešnosť v podobe získaných citácií. Ku každej množine záznamov sa zobrazí aj *h*-index.<sup>4</sup>

<sup>2</sup><http://ipscience.thomsonreuters.com>

<sup>3</sup>[http://images.webofknowledge.com/WOKRS57B4/help/WOS/hp\\_advanced\\_search.html](http://images.webofknowledge.com/WOKRS57B4/help/WOS/hp_advanced_search.html)

<sup>4</sup>[http://images.webofknowledge.com/WOKRS57B4/help/WOS/hp\\_citation\\_report.html](http://images.webofknowledge.com/WOKRS57B4/help/WOS/hp_citation_report.html)



### 1.2.2 Scopus

*Scopus* je citačný register európskeho vydavateľstva *Elsevier* so sídlom v Amsterdame, Holansko. Jedná sa o platenú službu, rovnako ako služba *Web of Science*. Bol spustený v novembri 2004, ale retrospektívne obsahuje záznamy od roku 1996. Scopus obsahuje viac záznamov hlavne z oblasti Európy. Okrem časopisov obsahuje aj zborníky, patenty a webové sídla. Aktualizuje sa denne. Podľa údajov z januára 2016 indexuje viac ako 21 500 titulov, ktoré zahŕňujú:

- viac než 21 500 recenzovaných časopisov (z toho 4 200 prístupných zdarma),
- viac než 360 obchodných časopisov,
- viac než 530 knižných edícií,
- viac než 7,2 milióna konferenčných príspevkov z 83 000 konferencií
- viac než 116 000 knižných titulov
- viac ako 27 miliónov patentových záznamov z piatich patentových úradov  
články v tlači (Articles-in-Press) z viac ako 5 000 časopisov

Databáza k januáru 2016 obsahuje viac než 60 miliónov záznamov v jadre, z toho:

- viac než 38 miliónov záznamov od roku 1996 (84% všetkých citácií),
- viac než 22 miliónov záznamov z obdobia 1823–1995 (staršie záznamy obsahujú len abstrakty bez citácií),
- okolo 3 milióna záznamov pribúda každým rokom (5 500 za deň).

V decembri 2015 bolo pridaných viac než 93 milióna citácií na viac než 5 miliónov článkov starších z pred roku 1996.<sup>5</sup>

Citačný register Scopus má veľmi dobre vyriešenú otázku identifikácie autorov a inštitúcií. Každý autor a inštitúcia má priradené všetky formáty mien resp. názvov, zároveň sú k dispozícii všetky dostupné štatistiky na základe indexovaných záznamov.

Napriek tomu, že pri možnosti spresnenia výsledkov sú pri jednotlivých možnostiach zobrazené aj počty záznamov (napríklad počty publikácií v jednotlivých rokoch), nie je možné tieto štatistiky prezentovať tak ako v databáze WoS. Funkcia **View citation overview** umožňuje analyzovať citácie označených záznamov. V prehľadnej tabuľke sa zobrazia publikácie a citácie v jednotlivých rokoch. **Author Evaluator** je nástroj zobrazujúci štatistiky a hodnoty indikátorov autorov na základe ich publikácií a získaných citácií. Koláčovým grafom sú prezentované časopisy, v ktorých autor publikoval, typy dokumentov a predmetové oblasti. Okrem toho je možné zobraziť počty publikácií a citácií v jednotlivých rokoch, všetkých spoluautorov a počty spoločných publikácií. Graficky sa zobrazí aj autorov *h*-index.

---

<sup>5</sup>[https://www.elsevier.com/\\_\\_data/assets/pdf\\_file/0007/69451/scopus\\_content\\_coverage\\_guide.pdf](https://www.elsevier.com/__data/assets/pdf_file/0007/69451/scopus_content_coverage_guide.pdf)

Funkcia **Affiliation details** zobrazí detailné informácie týkajúce sa konkrétnej inštitúcie. Ide o počet dokumentov, autorov, názvy časopisov, v ktorých títo autori publikujú. Okrem toho sú k dispozícii údaje o spolupráci s inými inštitúciami a koláčový graf zobrazujúci štruktúru predmetových oblastí na základe indexovaných publikácií.

**Journal Analyzer** je nástroj na hodnotenie a porovnávanie vedeckých časopisov. Podobnú funkciu má *Journal Citation Reports* v rámci *Web of Knowledge*. Na hodnotenie sú však zvolené iné indikátory. Ako alternatíva k tzv. impakt-faktoru sú uvedené SJR, SNIP, počty citácií v jednotlivých rokoch, percento necitovaných článkov a percento prehľadových článkov (reviews). Hodnoty sú prezentované graficky alebo v tabuľke. V grafe môžu byť na porovnanie zobrazené údaje o viacerých časopisoch.<sup>6</sup>

Databáza *Scopus* je silnou konkurenciou pre *Web of Science*, resp. *Web of Knowledge*. *Scopus* má širšie obsahové a teritoriálne zameranie, *WoS* zase dlhšiu tradíciu. Obidvom databázam nemožno uprieť snahu o rozšírenie svojho obsahu a funkcionality, z čoho má úžitok hlavne používateľ.

## 1.3 Citačné indikátory

Citačný indikátor je druh scientometrickej metódy na stanovenie „kvality“ vedeckých publikácií, vedeckých pracovníkov a vedeckých inštitúcií. Všetky indexy vychádzajú zo základných scientometrických parametrov: počet publikácií a množstvo ich citácií. Pri výpočte niektorých indexov zohľadňujú aj iné parametre (napr. vek pracovníkov).

Základným indikátorom je citačná frekvencia t.j. priemer počtu citácií istej skupiny publikácií v danom obore za určitý rok.<sup>7</sup>

### 1.3.1 Journal Impact Factor (IF)

Prvý citačný indikátor navrhol zakladateľ citačných registrov Eugen Garfield (1955) ako presnejší spôsob evaluácie autorov vedeckých článkov než v tej dobe používané počty publikácií a počty citácií.

V súčasnosti IF používa *Institute of Scientific Information* na každoročné hodnotenie vedeckých časopisov v rámci *Journal citation reports* (JCR). Impact Factor je priemerný počet citácií na články publikované danom časopise za posledné 2 roky.

Impact Factor pre rok 2016 možno matematicky vyjadriť vzťahom (1.8), kde  $a$  je celkový počet článkov, ktoré boli v danom časopise publikované v rokoch 2014–2015 a  $c$  je počet článkov publikovaných v danom časopise v rokoch 2014–2015, ktoré boli citované v publikáciách indexovaných v roku 2016.<sup>8</sup>

$$IF_{2016} = \frac{c}{a} \quad (1.8)$$

<sup>6</sup>[https://www.elsevier.com/\\_data/assets/pdf\\_file/0005/79196/scopus-quick-reference-guide.pdf](https://www.elsevier.com/_data/assets/pdf_file/0005/79196/scopus-quick-reference-guide.pdf)

<sup>7</sup><http://ipscience-help.thomsonreuters.com/incitesLiveESI/ESIGroup/fieldBaselines/citationRatesBaselines.html>

<sup>8</sup>Výsledný Impact Factor z roku 2016 môže byť publikovaný až v roku 2017, pretože ho nie je možné vypočítať skôr, než rok 2016 skončí.

JCR poskytuje IF za 5 ročné obdobie.<sup>9</sup>

### 1.3.2 Hirshov index (*h*-index)

Tento populárny citačný indikátor bol definovaný Jorge. E. Hirshom v roku 2005 ako číslo *h*, ktoré zodpovedá počtu najcitovanejších článkov daného autora, ktorých každá publikácia má aspoň *h* citácií. (Hirsh, 2005)

Pre lepšie pochopenie je vhodné uviesť príklad: Vedec A má 10 publikácií. Ak ich zoradíme podľa počtu citácií, potom prvá má 10 citácií, druhá má 8, tretia 5, štvrtá 4, piata 2 a ostatné nemajú žiadne citácie. Potom tento vedec má *h*-index 4, pretože štyri najcitovanejšie články (s počtami citácií 10, 8, 5 a 4) majú aspoň po 4 citácie. Tieto najcitovanejšie články sa označujú ako tzv. *h*-core.

Hlavným problémom *h*-indexu je necitlivosť na malý počet veľmi citovaných článkov (Napríklad ak porovnáme publikačnú činnosť vedca A s predchádzajúceho príkladu s vedcom B, ktorý má iba 5 publikácií so 108, 45, 12, 5 a 2 citáciami, jeho Hirshov index je rovnaký ako vedca A.

Sám Hirsh uviedol, že *H*-index nemožno použiť na porovnávanie autorov rôznych vedných disciplín.

### 1.3.3 Eggheov index (*g*-index)

Leo Egghe v roku 2006 publikoval indikátor *g*-index, ktorý má vyriešiť niektoré problémy *h*-indexu, najmä jeho necitlivosť k autorom, ktorí majú málo extrémne citovaných publikácií (Egghe 2006).

Eggheov *g*-index je definovaný ako číslo *g*, ktoré predstavuje počet najcitovanejších článkov konkrétneho autora, zostupne zoradený podľa počtu citácií, ktorého druhá mocnina je menšia alebo rovná súčtu všetkých citácií daných článkov.

Napríklad ak má vedec A desať publikácií, ktoré majú 6, 6, 5, 4, 2, 0, 0, 0, 0 a 0 citácií. Jeho *h*-index je 4 a *g*-index je 4. ( $6 + 6 + 5 + 4 = 21 \geq 4 \cdot 4 = 16$ ) Vedec B má šesť publikácií s 15, 10, 5, 4, 3 a 2 citáciami. Aj jeho *h*-index je tiež 4, ale *g*-index je 6. ( $15 + 10 + 5 + 4 + 3 + 2 = 39 \geq 6 \cdot 6 = 36$ )

Ako je z príkladu zrejmé,  $g\text{-index} \geq h\text{-index}$ . Keďže *g*-index berie do úvahy viacej citácií, ale stále je necitlivý ku autorom malého počtu extrémne citovaných článkov. Pokiaľ vedec má napr. 10 publikácií, ktoré sú spolu 300 citované, tak jeho maximálny *g*-index je 10 a zvyšných 200 citácií je ignorovaných. Z toho dôvodu sám Egghe navrhol umelo zvýšiť počet článkov na číslo *T*, ktorého druhá mocnina sa blíži ku celkovému počtu citácií (Egghe, 2008). Samozrejme manipulovanie s dátami nie je dobrá metóda a preto je potrebné použiť iný indikátor, ktorý je schopný zachytiť podobné prípady.

### 1.3.4 Zhangov *e*-index

Ako reakciu na malú citlivosť *h*-index a *g*-indexu pre autorov s malým množstvom veľmi citovaných prác Chun-Ting Zhang navrhol nový indikátor *e*-index. Zhang ho definoval ako

<sup>9</sup>[http://admin-apps.webofknowledge.com/JCR/help/h\\_impfact.htm](http://admin-apps.webofknowledge.com/JCR/help/h_impfact.htm)

číslo  $e$ , ktoré je druhou odmocninou rozdielu všetkých citácií  $h$ -core článkov a maximálnym počtom citácií, ktoré sú zahrnuté do  $h$ -indexu  $h^2$  (Zhang, 2009).

Napríklad, ak  $h$ -index akademického pracovníka je 10 a jeho publikácie v  $h$ -core majú spolu 200 citácií, tak jeho  $e$ -index bude 10, pretože ak odčítame teoretické minimum na dosiahnutie  $h$ -indexu 10, t.j. 100 citácií od skutočného počtu citácií  $h$ -core článkov 200, výsledok bude 100, z čoho druhá odmocnina je 10.

To znamená, že  $e$ -index možno použiť na odlíšenie dvoch vedcov s rovnakým  $h$ -indexom, ale rozdielnou citačnou frekvenciou.

### 1.3.5 Súčasný $h$ -index (Contemporary $h$ -index)

Autori citačných indikátorov si uvedomujú že vedecká literatúra starne. Všeobecne vedecká práca z pred 10 rokov má menší impakt, ako rok stará publikácia s rovnakým množstvom citácií. Sám Hirsh (2005) navrhol tzv.  $m$ -kvocient, čo nie je nič iné ako  $h$ -index podelený počtom rokov od vydania prvej práce daného vedca. Teda výrazne znevýhodňuje starších akademikov, bez ohľadu na to, či sú stále aktívni, a citovanosť ich najnovších publikácií.

Z toho dôvodu Sidiropoulos a kol. (2007) navrhli indikátor, ktorý zahrňuje vek jednotlivých článkov. Pomenovali ho Súčasný (contemporary)  $h$ -index  $h^c$ , ktorý definovali:

„Vedec má súčasný  $h$ -index  $h^c$  ak každý jeho článok z množiny  $N_p$  dosiahne skóre  $S^c(j) \geq h^c$  a ostatné články ( $N_p - h^c$ ) dosiahli skóre  $S^c(j) \geq h^c$ .“

Skóre  $S^c(j)$  je definované vzťahom (1.9), pričom  $Y(j)$  predstavuje rok, kedy bol článok  $j$  publikovaný a  $cit_j$  znamená jeho maximálny počet citácií.

$$S^c(j) = \gamma \cdot (Y(\text{teraz}) - Y(j) + 1)^{-\delta} \cdot cit_j \quad (1.9)$$

Pri nastavení  $\delta = -1$  sa dosiahne, že počet citácií daného článku je podelený jeho vekom v rokoch. Lenže, podľa autorov, podelením počtu citácií daného článku jeho vekom sa získajú príliš malé hodnoty skóre  $S^c(j)$  na dosiahnutie reprezentatívneho  $h$ -indexu. Preto autori zaviedli koeficient  $\gamma$ , ktorý podľa empirickej štúdie autorov je najvýhodnejšie nastaviť na  $\gamma = 4$ .

### 1.3.6 Citačná frekvencia váhovaná podľa veku (Age-weighted citation rate – AWCR) a AW-index

Vek vedeckej publikácie (rozdiel medzi dnešným rokom a rokom vydania daného článku) sa považuje za jeden z faktorov, ktorý definuje impakt článku. Na jeho kvantifikáciu je nutné započítať tento parameter do výpočtu. Jednoduchým delením počtu citácií danej publikácie jej vekom váhujeme citačnú frekvenciu podľa veku (publikácie).

Jin a kol (2007) vytvorili indikátor, ktorý váhuje citačnú frekvenciu  $h$ -core článkov podľa veku. Nazvali ho AR-index.

Jeho matematická definícia je vyjadrená vzťahom (1.10), pričom  $h$  je Hirshov index autora,  $cit_j$  je množstvo citácií  $j$ -teho najcitovanejšieho článku a  $a_j$  je počet rokov od publikácie  $j$ -teho článku.

$$AR = \sqrt{\sum_{j=1}^h \frac{cit_j}{a_j}} \quad (1.10)$$

Autori programu Publish or Perish vytvorili *AW-index* – modifikáciu *AR-indexu*. Na rozdiel od *AR-indexu*. *AW-index* berie do úvahy všetky publikácie, nie len začlenené v *h-core*.

### 1.3.7 Individuálny (Individual) *h-index*

Batista a kol. (2006) navrhli nový index –  $h_I$ , ktorý by bol multidisciplinárny na rozdiel od *h-indexu*. Kdeže jeden z hlavných rozdielov medzi vedeckými disciplínami je množstvo vedcov, ktorí pracujú v danej disciplíne, Batista a kol. definovali  $h_I$ -index ako podiel *h-indexu* s priemerným počtom autorov *h-core* článkov.

Matematicky ho definovali vzťahom (1.11), kde  $h$  je Hirshov index a  $\langle N_a \rangle = N_a^{(T)} / h$ , pričom  $N_a^{(T)}$  je celkový počet autorov (vrátane opakovaní) *h-core* článkov.

$$h_I = \frac{h}{\langle N_a \rangle} = \frac{h^2}{N_a^{(T)}} \quad (1.11)$$

### 1.3.8 Individuálny (Individual) *h-index* $h_{I,norm}$

Autori programu *Publish or Perish* spravili modifikáciu individuálneho *h-indexu*  $h_I$ . Tento index pomenovali  $h_{I,norm}$ . Na rozdiel od  $h_I$ , jednoducho *h-index* delili počtom všetkých spoluautorov *h-core* článkov. Pri výpočte  $h_{I,norm}$  sa počet citácií jednotlivých článkov podelí počtom autorov daného článku. A potom sa vypočíta Hirshov index z už takto znormalizovaných publikácií.

### 1.3.9 Multi-autorský *h-index*

Michael Schreiber (2008) popísal nový indikátor, ktorý zahrňuje spoluautorstvo –  $h_m$ -index. Schreiber ho odvodil od individuálneho *h-indexu*  $h_I$  s tým rozdielom, že počtom autorov je delený rang dokumentu, nie počet citácií, ako v  $h_I$  a z toho sa vypočíta *h-index*.

Matematicky je definovaný vzťahom (1.12), pričom  $r$  je tzv. rang publikácie v zostupnom zoradení podľa počtu citácií,  $c(r)$  je počet citácií článku  $r$  a  $r_{eff}(r)$  je efektívny rang článku  $r$ .

$$h_m = \max_r (r_{eff}(r) \leq c(r)) \quad (1.12)$$

Efektívny rang článku  $r$  je definovaný vzťahom (1.13), kde  $a(r')$  je počet autorov publikácie  $r'$ .

$$r_{eff}(r) = \sum_{r'=1}^r \frac{1}{a(r')} \quad (1.13)$$

## Kapitola 2

# Druhá kapitola s matematikou $\int f(x) dx$ v názvu

[illegible]

$$\int f(x) \mathrm{d}x \quad (\text{rovnice})$$

odkaz na rovnici s „tagem“ (**rovnice**) – pokud nechcete vzorci přidělit číslo, ale nějaký vlastní symbol, používejte „hvězdičkovaná“ prostředí, tj. např. `equation*`

$$\iint f(x) \, dx \quad (2.1)$$

odkaz na druhy vzorec (3.1)

## 2.1 Podkapitola

# Kapitola 3

# Druhá kapitola s matematikou $\int f(x) dx$ v názvu

[illegible]

$$\int f(x) \mathrm{d}x \quad (\text{rovnice})$$

odkaz na rovnici s „tagem“ (**rovnice**) – pokud nechcete vzorci přidělit číslo, ale nějaký vlastní symbol, používejte „hvězdičkovaná“ prostředí, tj. např. equation\*

$$\iint f(x) \, dx \quad (3.1)$$

odkaz na druhy vzorec (3.1)

### 3.1 Podkapitola

# Záver

Scientometria je hodnotenie vedy, vedeckých publikácií, vedeckých pracovníkov a vedeckého pokroku použitím matematických, štatistických metód. Hlavným aspektom, podľa ktorého sa hodnotia vedecké práce sú citácie, t.j. referencie na iné publikácie, ktoré autor použil, alebo chce na ne upozorniť. Všeobecne je brané, že publikácia, ktorú cituje viacej iných vedeckých článkov má väčší impakt. To znamená, že práca je populárna, pretože je kvalitná a prínosná pre vedecký pokrok.

Ďalším spôsobom hodnotením vedy je použitie ekonomických aspektov. Počet a hodnota grantov, ktoré daný pracovník, či inštitúcia dosiahla, alebo hodnota praktického uplatnenia konkrétnych poznatkov. Bohužiaľ tento spôsob hodnotenia nemôže byť všeobecný a ani spravodlivý. Pretože cieľom vedy nie je vytvoriť prospešný a ekonomicky výhodný produkt, ale posunúť ľudské poznanie. Väčšina výskumu patrí do tzv. základného výskumu, u ktorého sa neočakáva možnosť aplikácie nadobudnutých poznatkov do praxe. Mnohé z nich nie sú doteraz aplikovateľné a niektoré našli praktické uplatnenie až po uplynutí niekoľko storočí (napr. matematické modely umelej inteligencie). V neposlednom rade predmetom výskumu je overenie hypotézy. Vedec by mal očakávať, že výsledok výskumu bude vyvrátenie hypotézy a automaticky neprinesie ekonomický úžitok, ale iba pokrok. Každý vedec musí mať na pamäti, že aj negatívny výsledok je výsledok.



# Príloha

**Sem môžete pridať prílohu. Ak chcete „Prílohy“, tak upravte definíciu záhlavia v súbore `ucm.thesis.sty`, viď príkaz `\HlavickaPriloha`.**

# Zoznam použitej literatúry

BELLIS, N. D., 2009. *Bibliometrics and citation analysis : from the Science citation index to cybermetrics*. The Scarecrow Press, Inc., 2009. ISBN 9780810867130.

ONDRIŠOVÁ, M., 2011. *Bibliometria*. STIMUL, 2011. ISBN 9788081270352.

PRITCHARD, A., 1969. Statistical bibliography or bibliometrics? In *Journal of Documentation*. 1969, 25, s. 348–349. ISSN 0022-0418.

