

# Hat előadás a tudományometriáról

*Schubert András*



Budapest, 2022

# Hat előadás a tudományometriáról

*Schubert András*



**ELTE**  
EÖTVÖS LORÁND  
TUDOMÁNYEGYETEM

**ELTE BTK**

**Könyvtár- és Információtudományi Intézet**

**Budapest, 2022**

**ISBN 978-963-489-432-2**

**DOI 10.21862/978-963-489-432-2**

## 0. Bevezetés

Ez az írás nem kézikönyv, nem is tankönyv, talán egyáltalán nem is könyv. Egy hat előadásból álló egyetemi kurzus (amelyet az ELTE BTK Könyvtár- és Információtudományi Intézetében tartottam) hevenyészett vázlatából próbáltam egy kissé rendezettebb formájú közreadható művet kerekíteni.

Serkentően és egyben nyomasztóan hathat, hogy kutatói pályám kezdetének olyan meghatározó könyveivel kerülhetek formai rokonságba, mint Truesdell [1966] vagy Gantmacher [1975] előadaskötetei, vagy éppenséggel a tudománymetriának de Solla Price 1962-es Pegram Lectures előadásaiból készült mérföldköve [Price, 1963].

Az előadások alapfokú bevezetést próbálnak nyújtani a tudománymetria történetébe, művelésébe és alkalmazásába. A rendelkezésre álló idő, illetve terjedelem, valamint – mi tagadás – saját érdeklődésem és ismereteim korlátai miatt a felölelt témákat bizonyos mértékben önkényesen választottam ki. Ha úgy tetszik, adhatjuk a munkának ezt a címet is: „Ízelítő a tudománymetriából”. A gasztronómiai metafora nem egészen esetleges; a későbbiekben még vissza-visszatérünk rá.

### KITÉRŐ

A szakácskönyv hasonlat nem újdonság a tudományban. A számítástudományban a „cookbook” az elmúlt évtizedekben sajátos jelentésre tett szert: a programozási nyelvek kis programrészletek („snippet”), vagyis „receptek” segítségével történő megismertetését nevezik így. A programozási segédkönyvek két vezető kiadója, a Packt Publishing és az O’Reilly Media ezernél is több ilyen című könyvet adott ki az utóbbi években.

Ez a metafora a könyvtár- és információtudomány területén sem ismeretlen. A tudományos publikációkkal foglalkozó egyik vezető blog, a Scholarly Kitchen [scholarlykitchen.sspnet.org] szerkesztői „séfeknek” nevezik magukat, és a blog jelmondata: „What’s Hot and Cooking In Scholarly Publishing”, szabadon fordítva: Mi fő a tudományos publikálás fazekában?

Az „Információs műveltség szakácskönyve” [Secker et al., 2007] a siker „hozzávalóit, receptjeit és ötleteit” ígéri alcímében.

Remélem, a következőkben összegyűjtött falatok nem fekszik meg az olvasó gyomrárt!

# 1. A tudománymetria helye és szerepe

## 1.1. Miről is beszélünk?

„Mi a tudománymetria?” Ez a kérdés annyi tudományos és kevésbé tudományos cikk, könyvfejezet, brosúra címében vagy alcímében szerepel, hogy az embernek önkénytelenül kétségei támadnak, hogy tudja-e bárki is a kérdésre a választ. Természetesen, ha utánanézzünk a „Mi a fizika?” vagy „Mi a kémia?” kérdéseknek, akkor rájöhethetünk, hogy ebben a vonatkozásban egyik tudománynak sincsen mit a másik szemére vetni.

Ennek kapcsán persze elgondolkozhatunk azon, hogy van-e ezekben az esetekben a pontos definícióknak jelentősége, és ha van, akkor mi az. Engedje meg az Olvasó, hogy a Szerző megossza vele az ezzel kapcsolatos kissé sarkos véleményét: a pontos definícióknak két esetben van nagy jelentősége, (1) a matematikában, (2) jogi – főként pénzügyi – kérdésekben.

Mindezek előrebocsátásával próbáljuk azért körüljárni a tudománymetria meghatározását.

A tudománymetria egy lehetséges „hivatalos” definíciójának tekinthetjük például az erre a témára szakosodott nemzetközi folyóirat „önmeghatározását” (1. ábra):



1. ábra A *Scientometrics* című folyóirat fedőlapjának részlete

Eszerint a tudománymetria a tudományos kommunikáció és a tudománypolitika minden mennyiségi vonatkozását foglalja magába.

A folyóirat honlapjának részletesebb meghatározása szerint:

„Scientometrics is concerned with the quantitative features and characteristics of science and scientific research. Emphasis is placed on investigations in which the development and mechanism of science are studied by statistical mathematical methods.”

[[www.springer.com/journal/11192](http://www.springer.com/journal/11192)]

Vagyis: A tudománymetria a tudomány és a tudományos kutatás mennyiségi tulajdonságaival és jellemzőivel foglalkozik. Különleges hangsúlyt kapnak azok a vizsgálatok, amelyek a tudomány fejlődését és mechanizmusait statisztikai, matematikai módszerekkel tanulmányozzák.

A Wikipédia szerint: „Scientometrics is the field of study which concerns itself with measuring and analysing scholarly literature” [[en.wikipedia.org/wiki/Scientometrics](http://en.wikipedia.org/wiki/Scientometrics)]. Tehát a tudománymetria a kutatás olyan területe, amely a tudományos irodalom mérésével és elemzésével foglalkozik.

A magyar Wikipédián „tudománymetria” címszó nem található; a keresés a „A tudományos teljesítmény mérése” címszóhoz vezet.

Ez az észrevétel egybevág Soós Sándor írásának bevezetésével: „Ha egy, a kutatói társadalomban végzett közvéleménykutatás keretében arra kérnénk a megkérdezetteket, definiálják a tudománymetria fogalmát, a sor elején valószínűleg olyan válaszok állnának, amelyek szerint a tudománymetria a kutatói teljesítményértékelés (főként adminisztratív) eszköze, célja a tudomány szereplőinek minőség-ellenőrzése, megrendelője (haszonélvezője) pedig a tudomány szakpolitikusa. Ha ugyanakkor elővesszük a Scientometrics nevű, sok évtizedre visszatekintő nemzetközi – a terület hazai szaktekintélyei

által indított és szerkesztett – szaklapot, annak borítóján a következő meghatározással szembesülhetünk: «a tudományos kommunikáció kvantitatív vizsgálata».

A „hivatalos” definíciók láthatóan szerteágazóak, némiképpen talán ellentmondóak is. Ha azonban nem matematikusként vagy jogászként tekintjük a kérdést, ettől nem kell feltétlenül visszariadnunk. Egy szakácskönyv sem lesz használhatatlan azért, mert nem találjuk meg benne a leves pontos definícióját.

Ha „nem hivatalosan” tesszük fel a kérdést – Mi a tudománymetria? – a válasz ez lehet: sok minden. Nézzünk néhány lehetséges és néhány alig lehetséges példát

### ***1.2. A tudománymetria, mint tudomány***

A tudománymetriát művelőinek és ismerőinek nagy része a tudomány, ezen belül a társadalomtudomány egyik területnek tekinti.

A tudományterületek körülhatárolása magában is egy kutatási terület, sőt nem is csak egy, hiszen többféle keretben is megkísérelhető.

Az egyes tudományterületek – egyebek között – különböznek

- a tárgyukban (ismeretelméleti megközelítés);
- a módszerükben (módszertani megközelítés);
- a célközönségükben (szociológiai megközelítés).

Ha ezek szerint a szempontok szerint próbáljuk meg meghatározni a tudománymetriát, akkor a következő jellemzéshez juthatunk:

- A tudománymetria tárgya a tudományos kutatási tevékenység, különös tekintettel a tudományos ismeretek kommunikációjára.
- A tudománymetria módszere bármiféle kvantitatív vizsgálati módszer, tehát voltaképpen a matematika és a statisztika bármilyen eszköze.
- A tudománymetria célközönsége korántsem egységes. Ide sorolhatjuk
  - (i) a különféle szaktudományok művelőit, akiknek a számára a tudománymetria az önértékelés egy lehetséges módszere;
  - (ii) a tudománymetria és rokon információtudományok művelőit, akik természetesen folyamatosan nyomon követik a szakma legújabb fejleményeit, különösképpen a módszertani újdonságokat;
  - (iii) a tudománypolitika művelőit, akiket természetesen a tudománymetriai módszerek gyakorlati alkalmazása, a rangsorolási, értékelési kérdések foglalkoztatnak a leginkább.

#### KITÉRŐ

A könyvtárosokat – talán nem bántódnak meg érte – a (ii) kategóriába sorolnám, azok közé, akik a tudománymetriát – legalábbis potenciálisan – művelik. Az a tapasztalatom, hogy a kutatóintézmények, egyetemek a ma már szinte mindenütt nélkülözhetetlen tudománymetriai tevékenységet a könyvtárakhoz rendelik – még ha esetleg maga a könyvtár, illetve a könyvtárosok nem értenek is ezzel teljesen egyet. Ennek megfelelően szép számmal található *Tudománymetria (vagy bibliometria) könyvtárosoknak* című egyetemi kurzusok, előadásanyagok, sőt teljes konferenciák is az interneten. A tudománymetria mindenképpen hozzátartozik a korszerű könyvtárosi munka eszköztárához – különösképpen a kutatóintézmények, egyetemek könyvtáraiban.



A három szempontnak (tárgy, módszer, célközönség) együttesen kell érvényesülni; ha egy kutatásnak akár a tárgya, akár a módszere, akár a célközönsége eltér a megjelöltektől, akkor azt aligha lehet a tudománymetria területéhez sorolni. Ez nemcsak afféle elvi okoskodás: a Scientometrics című folyóirathoz beküldött sokszáz kézirat között mindig akad egy-két tucat, amelynek szerzője láthatóan nem a fentiek szerint értelmezi a folyóirat témáját, és ezért – a munka esetleges egyéb kétségtelen értékei ellenére – vissza kell utasítanunk a kéziratot.

A tudománymetriának, mint kutatási területnek kialakultak a saját alterületei. A tudományos kutatás bármely fő- és alterületéhez hasonlóan ezeknek a körülhatárolása korántsem egyértelmű és általánosan elfogadott. Egy lehetséges felosztás például a következő:

- strukturális tudománymetria;
  - dinamikai tudománymetria;
  - értékelő tudománymetria
- A strukturális tudománymetria a tudományos kutatás résztvevői (ebbe a legtágabb értelemben beleértve a kutatókat, intézményeket, folyóiratokat, kiadókat, országokat stb.) között fennálló szerkezeti kapcsolatokat vizsgálja. A manapság már „hagyományosnak” tekinthető többváltozós statisztikai módszerek (faktoranalízis, klaszterelemzés stb.) mellett az utóbbi évtizedekben óriási fejlődésen keresztülment hálózatelemzés a módszerek bőséges tárházát kínálja. A kutatók számára az igazi kihívást valójában nem az elemzések technikai végrehajtása jelenti, mert ezekhez a szabadon hozzáférhető és a kommerciális statisztikai programcsomagok „konyhakész” rutinok tömegét kínálják. Az igazi nehézség a legmegfelelőbb módszer kiválasztásában, és még inkább a kapott eredmények értelmezésében van. Ez utóbbihoz nem elegendő a felhasznált módszer bármilyen alapos ismerete,

hanem szükség van a vizsgálat tárgyát képező kutatási területnek és résztvevőinek beható ismeretére is.

➤ A dinamikai tudománymetria a kutatási folyamatok, illetve az azokat jellemző mennyiségi mutatószámok térbeli-időbeli változásait vizsgálja. Klasszikus példaként említhetjük a publikációk számának exponenciális és logisztikus modelljeit, vagy a kutatói migráció nyomon követését – ez utóbbit akár földrajzi, akár tudományterületi értelemben értjük is.

➤ Az értékelő (evaluatív) tudománymetria, ahogyan már céloztunk rá, a tudománymetria leginkább szem előtt lévő, és egyben legvitatottabb területe. A „nagyközönség”, de még a tudománymetriai kutatásokat aktívan művelők számára is a rangsorolási, értékelési lehetőségek adják a tudománymetriai tevékenység létjogosultságát. Az „impakt faktor”, a „h-index”, az egyetemi rangsorok nagy teret kapnak a tudományos közéletben, sőt a napisajtóban is. Kétségtelen, hogy az értékeléseken keresztül jelenik meg a tudománymetriai kutatások társadalmi hatása („societal impact”). Minden más valójában csak a „bennfentesek” számára érdekes, de éppen ez a háttértudás az, amely az értékeléseknek megbízhatóságot és relevanciát biztosít. Ahhoz szoktam ezt hasonlítani, mint ahogy a meteorológia tudományának értékét a nagyközönség az időjárásjelentések pontosságán méri le, ami valójában a meteorológusok tevékenységének csak egy rendkívül szűk, és korántsem a legnagyobb kihívást jelentő szelete.

A tudománymetria felsorolt három alterülete nem egymást kizáró. A kutatói hálózatok dinamikája, a nagyon gyorsan elterjedő, „forró” témák mutatószámain alapuló rangsorok, a hálózatokban való elhelyezkedés – a centralitás – értékmérő szerepe az említett részterületek átfedésének példái.

## KITÉRŐ

A strukturális-értékelő tudományometriai mutatószámok egy nevezetes példája az Erdős-szám.

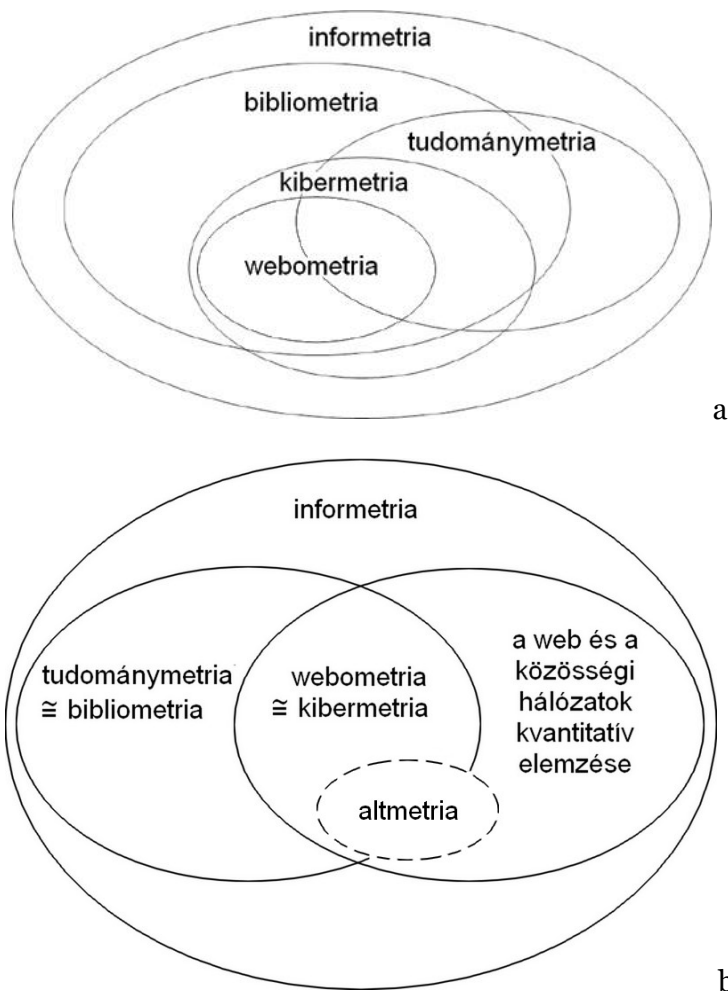
Erdős Pál (1913–1996) a XX. század egyik legkiemelkedőbb matematikusa volt. A matematikusok között szokatlan módon cikkeinek jelentős részét társszerzőkkel írta. Munkássága során összesen 511 szerzővel írt közösen cikket. Közéjük tartozni a matematikus-társadalomban nagy elismerésnek számít. Az Erdős Páltól való „társszerzői távolság”, az Erdős-szám, számontartott értékmérő.

Definíció: Erdős Pál Erdős-száma  $o$ . Egy szerző Erdős-száma  $n$ , ha társszerzőinek Erdős-száma közül a legkisebb érték  $n-1$ .

A Nobel-díjasoknak és a tudományos közösség kiemelkedő személyiségeinek az Erdős-száma általában 2–4 között van (Einstein: 2, Neumann: 3, Pauling: 4, Eccles: 3, Gates: 4).

A tudománymetria mellett az elmúlt évtizedekben több olyan rokon tudományterület jelent meg, amelyek tárgyukban, módszereikben, célközönségükben részben átfednek vele és egymással: bibliometria, informetria, kibernetika, webometria, altmetria stb. Szükség van-e ezeknek a területeknek a pontos megkülönböztetésére, meghatározására? Véleményemet a pontos definíciók szükségességéről fentebb elmondtam. Maradjunk annyiban, hogy egyeseknek van igénye ilyen meghatározásokra, másoknak kevésbé.

A 2. ábrán bemutatok két olyan sémát, amely ezeket a területeket valahogyan rendszerezni próbálja.



2. ábra A tudománymetria és a vele rokon területek elhelyezkedése  
a: Björneborn & Ingwersen [2004] szerint; b: Schubert [2015] szerint

### **1.3. A tudománymetria, mint gyakorlati módszer**

Soós Sándor fentebb idézett megállapítása szerint a tudománymetriát sokan a tudománypolitika „adminisztratív eszközének” tekintik. Nyilván nem is alaptalanul, mert az egyik arculata éppen ez. A tudománymetriának mind a három előbb felsorolt részterülete a módszertani alapvetésen kívül gyakorlati módszereket is kínál a kutatómenedzsment vagy éppen a tudománypolitika számára.

➤ A strukturális tudománymetria nagy szolgálatot tehet a kutatási stratégiák kialakításában a tudományterületi és együttműködési térképek készítésével. Mind az egyéni kutatók, mind az intézmények, és különösképpen a folyóiratok számára rendkívül hasznosak a hálózati struktúrákon alapuló ajánlórendszerek. Ezek segíthetnek együttműködő partnereket, az elért eredményeket a leghatékonyabban közlő folyóiratokat, vagy kéziratokhoz bírálókat találni.

➤ A dinamikai tudománymetria a térbeli és időbeli változások követésével és lehetőség szerinti előrejelzésével segítheti az eredményes kutatást. A trendelemzések módszertana jól ismert a gazdasági-pénzügyi világból. Ezeknek a módszereknek az adaptálása a tudományos kutatás területére rendkívül nagy körültekintést igényel. Enélkül a kapott eredmények könnyen érdektelenek, vagy adott esetben éppenséggel félrevezetőek lehetnek.

➤ Az értékelő tudománymetria *par excellence* a tudománymetria leginkább gyakorlati célokat szolgáló ága. Egyszersmind a legvitatottabb is. Ezt egyáltalában nem lehet csodálni. Sok emberből a pusztá számbavétel, nyilvántartás is ellenérzéseket vált ki (nyilván sok-sok kedvezőtlen tapasztalat vezetett ide), a minősítésre, értékelésre ez hatványozottan igaz. Természetesen az ellenérzések gyökere nem az eszközökben, hanem a használat módjában, vagy talán még inkább az eszközöket használó emberek szándékaiban keresendő. Ezeknek a szándékoknak a felderítése vagy módosítása azonban a legkevésbé sem a tudománymetria kompetenciája vagy feladata.

Elkerülhetetlen, és miért is kellene elkerülni, hogy a gyakorlatban használható módszereket valaki, valahol, valamikor használni is fogja. Az értékelő tudománymetria módszereit igenis fogják kutatók, kutatócsoportok, intézmények, folyóiratok stb. értékelésére használni. Ezek az értékelések

embereket érintő döntéseket fognak befolyásolni. A tudománymetria művelőinek a felelőssége, hogy e módszerek használatának feltételeit, korlátait, az eredmények értelmezésének lehetőségét a döntéshozók tudomására hozzák. Maga a döntés azonban a döntéshozók felelőssége, és a leghatározottabban tiltakozni kell az ellen, hogy ezt a felelősséget megpróbálják áthárítani a „metriára”, a „mutatószámokra” és ezen keresztül a tudománymetria művelőire. Az orvosi vizsgálatok számszerű vagy képi leletei nélkülözhetetlenek az orvos számára, de a döntés, a diagnózis és a terápia meghatározása az orvos felelőssége, és azt nem háríthatja át a laboratóriumi vegyészekre vagy a képalkotó vizsgálatok készítőire.

#### KITÉRŐ

Az orvosi döntések megkönnyítését szolgálják azok a „irányelvek” (*guidelines*), amelyek a diagnosztikai és terápiás eljárásokra ajánlásokat, alkalmanként előírásokat fogalmaznak meg. Ezek a protokollok a „bizonyítékokon alapuló” (*evidence based*) orvoslás paradigmájának a keretében születnek körültekintő statisztikai elemzések alapján.



Természetesen nagy igény lenne a tudománymetria gyakorlati alkalmazását megkönnyítő hasonló irányelvekre, és történtek erre vonatkozó próbálkozások. Példaként tekinthetjük a DORA deklarációt, amelyről később majd még egy kicsit több szó esik. A meglévő példák azonban inkább intőek, mint követendőek. Ennek a fő oka az, hogy a tudománymetria alkalmazásában mind a célokat, mind az eszközöket tekintve sokkal kisebb az egyetértés, mint az orvoslásban. Ezt pedig sem elvi megfontolásokkal, sem hatalmi szóval nem lehet helyettesíteni.

A gyakorlati módszereket az elvi fejtegetéseknél bizonyára hasznosabban illusztrálja néhány gyakorlati példa. Ez akkor is igaz (vagy akkor még inkább igaz), ha a példák egyike-másika nem igazán követendő.

➤ A Web of Science-nek kezdetektől fogva működtet egy hasznos ajánlórendszert „hasznló tételek” (Related Records) néven. A rendszer, amelyet Kessler [1963] elvi alapvetése nyomán George Vladutz, az Institute for Scientific Information munkatársa dolgozott ki [Vladutz & Cook, 1984], az adatbázisban nyilvántartott minden publikációhoz megkeresi a „leghasznlóbb” tételt. A hasonlóságot a bibliográfiai csatolás módszere alapján, a cikkek hivatkozási listájában szereplő közös tételek aránya alapján számítják ki.

Hasonló ajánlórendszerek működnek a legtöbb bibliográfiai adatbázisban, mint pl. a Scopusban vagy a Google Scholarban. A Web of Science-től eltérően többnyire nemcsak a közös hivatkozásokon alapuló kapcsolatokat, hanem egyéb tényezőket (közös címszavak, kulcsszavak stb.) is figyelembe vesznek a hasonlóság megállapításánál. A pontos algoritmus általában nem nyilvános.

Az 1. ábrában hivatkozott Björneborn & Ingwersen [2004] cikkhez leghasznlóbb cikkeket mutatjuk meg a Web of Science, a Scopus és a Google Scholar alapján a 3. ábrán.

<input type="checkbox"/> 1 	<b>Webometrics</b> <a href="#">Thelwall, M; Vaughan, L and Bjerneborn, L</a> 2005   <a href="#">ANNUAL REVIEW OF INFORMATION SCIENCE AND TECHNOLOGY</a> 39 , pp.81-135	<b>134</b> Citations <hr/> <b>240</b> References ( 50 shared )
<a href="#">Full Text at Publisher</a> ***		<a href="#">Related records</a>
<input type="checkbox"/> 2 	<b>Informetrics at the beginning of the 21st century - A review</b> <a href="#">Bar-Ilan, J</a> 2008   <a href="#">JOURNAL OF INFORMETRICS</a> 2 (1) , pp.1-52 This paper reviews developments in informetrics between 2000 and 2006. At the beginning of the 21st century we witness considerable growth in webometrics, mapping and visualization and ope: ... <a href="#">Show more</a>	<b>240</b> Citations <hr/> <b>609</b> References ( 20 shared )
<a href="#">Full Text at Publisher</a> ***		<a href="#">Related records</a>
<input type="checkbox"/> 3	<b>Mathematical models for academic webs: Linear relationship or non-linear power law?</b> <a href="#">Payne, N and Thelwall, M</a> Dec 2005   <a href="#">INFORMATION PROCESSING &amp; MANAGEMENT</a> 41 (6) , pp.1495-1510 Previous studies of academic web interlinking have tended to hypothesise that the relationship between the research of a university and links to or from its web site should follow a linear trend, yet 1 ... <a href="#">Show more</a>	<b>2</b> Citations <hr/> <b>65</b> References ( 16 shared )
<a href="#">Full Text at Publisher</a> ***		<a href="#">Related records</a>

(a)

## Informetric analyses on the world wide web: methodological approaches to 'webometrics'

TC Almind, [P Ingwersen](#) - *Journal of documentation*, 1997 - emerald.com


This article introduces the application of informetric methods to the World Wide Web (WWW), also called Webometrics. A case study presents a workable method for general informetric analyses of the WWW. In detail, the paper describes a number of specific informetric ...

☆  Cited by 1244 Related articles All 14 versions

### Webometrics

[M Thelwall, L Vaughan...](#) - *Annual review of ...*, 2005 - Wiley Online Library


Webometrics, the quantitative study of Web-related phenomena, emerged from the realization that methods originally designed for bibliometric analysis of scientific journal article citation patterns could be applied to the Web, with commercial search engines ...

☆  Cited by 456 Related articles All 12 versions

### The calculation of web impact factors

[P Ingwersen](#) - *Journal of documentation*, 1998 - emerald.com

This case study reports the investigations into the feasibility and reliability of calculating impact factors for web sites, called Web Impact Factors (Web-IF). The study analyses a selection of seven small and medium scale national and four large web domains as well as ...

☆  Cited by 944 Related articles All 14 versions

(b)



	Document title	Authors	Year	Source	Cited by
<input type="checkbox"/> 1	Webometrics	Thelwall, M., Vaughan, L., Björneborn, L.	2005	Annual Review of Information Science and Technology 39, pp. 81-135	169
	<a href="#">View at Publisher</a> <a href="#">Related documents</a>				
<input type="checkbox"/> 2	Limits and feasibility of Cositation method on the web. An experiment on the French speaking web	Prime-Claverie, C., Belgbeder, M., Lafouge, T.	2005	Proceedings of ISSI 2005: 10th International Conference of the International Society for Scientometrics and Informetrics 1, pp. 78-86	2
	<a href="#">Related documents</a>				
<input type="checkbox"/> 3	'Mini small worlds' of shortest link paths crossing domain boundaries in an academic Web space	Björneborn, L.	2006	Scientometrics 68(3), pp. 395-414	20
	<a href="#">View abstract</a> <a href="#">View at Publisher</a> <a href="#">Related documents</a>				

(c)

3. ábra Björneborn & Ingwersen [2004] cikkéhez leghasonlóbb tételek (a) a Web of Science “Related Records”, (b) a Google Scholar “Related articles” és (c) a Scopus “Related documents” ajánlórendszere szerint.

A “legközelebbi” tételek láthatóan eltérnek, de megállapíthatjuk, hogy a találatok minden esetben relevánsak; valóban közeli rokonságban vannak az „anyacikk” témájával.

➤ Ha elvégzünk egy egyszerű Google keresést a „trendelemzések bibliometriai módszerekkel” témában, megdöbbentő találatszámot láthatunk (4. ábra). Ne higgyünk a látszatnak! Ha lelapozunk a találati lista aljára, a következő megjegyzést találjuk: „In order to show you the most relevant results, we have omitted some entries very similar to the 130 already displayed.” Valójában tehát százegynéhány érdemi találatot duzzaszt fel a Google sajátos számlálási rendszere többmillióra. Intő példa a Google-alapú statisztikák megbízhatatlanságára.

Persze a százas nagyságrendű találati szám sem kevés; a szóban forgó témának jelentős irodalma van. Jegyezzük azért meg, hogy bár a találatok jelentős része tudományos publikációkra utal, valójában ezek a munkák inkább „ipari termékek”, mint kutatási eredmények.

The image shows a Google search results page for the query "bibliometric trend analysis". The search bar at the top contains the text "bibliometric trend analysis" with a search icon on the right. Below the search bar, there are navigation tabs for "All", "Images", "Videos", "News", "Shopping", and "More", along with a "Tools" link. The search results are displayed below, starting with "About 3,070,000 results (0,64 seconds)". The first result is titled "Scholarly articles for bibliometric trend analysis" and lists several articles with their titles and citation counts. The second result is a link to a full-text article from SPIE Digital Library, titled "Bibliometric trend analysis on global ... - SPIE Digital Library", by S Chen Sr, dated 2018, with 1 citation. The third result is from BMC Complement Med Therapies, titled "Bibliometric analysis of trends and issues in traditional ...", by L Huang, dated 2020, with 4 citations. The fourth result is from Mattioli1885journals.com, titled "Bibliometric trend analysis in a decade of European ...", dated July 1, 2021. The fifth result is from the Journal of Big Data, titled "A bibliometric approach to tracking big data research trends", by A Kalantari, dated 2017, with 60 citations. The sixth result is from Frontiers, titled "Bibliometric Analysis of Global Research Trends on ... - Frontiers", dated April 22, 2021. The seventh result is from PubMed, titled "Bibliometric trend analysis in a decade of European ... - PubMed", by O Mani, dated 2021.

Search results for "bibliometric trend analysis":

- About 3,070,000 results (0,64 seconds)
- Scholarly articles for **bibliometric trend analysis**
  - Bibliometric trend analysis** on global graphene ... - Lv - Cited by 104
  - ... : **Bibliometric analysis** as a tool for mapping of **trends** ... - Li - Cited by 301
  - Bibliometric analysis**-a new business area for ... - Ball - Cited by 121
- <https://www.spiedigitallibrary.org> > 12.2505675.full > **Bibliometric trend analysis on global ... - SPIE Digital Library**  
by S Chen Sr · 2018 · Cited by 1 — **Bibliometrics** is statistical **analysis** of written publications, such as books or articles. **Bibliometric** methods are frequently used in the field ...
- <https://bmccomplementmedtherapies.biomedcentral.com> > ... > **Bibliometric analysis of trends and issues in traditional ...**  
by L Huang · 2020 · Cited by 4 — This study explored the **trends** and issues relating to the application of traditional medicine in stroke research. A **bibliometric** search was ...
- <https://www.mattioli1885journals.com> > article > view > **Bibliometric trend analysis in a decade of European ...**  
Jul 1, 2021 — Purpose The purpose of this **bibliometric** study was to summarize European orthopedic literature produced by EFORT memberships between 2009 ...
- <https://journalofbigdata.springeropen.com> > articles > **A bibliometric approach to tracking big data research trends**  
by A Kalantari · 2017 · Cited by 60 — The methodology is based on **bibliometric** techniques which permit a robust **analysis** of "Big Data Research" publications at different levels. The ...
- <https://www.frontiersin.org> > fphar.2021.646626 > full > **Bibliometric Analysis of Global Research Trends on ... - Frontiers**  
Apr 22, 2021 — **Bibliometrics** is the use of mathematical–statistical methods for summarizing scientific activities in a research field and further identifying ...
- <https://pubmed.ncbi.nlm.nih.gov> > ... > **Bibliometric trend analysis in a decade of European ... - PubMed**  
by O Mani · 2021 — Purpose The purpose of this **bibliometric** study was to summarize European orthopedic literature produced by EFORT memberships between 2009 ...

#### 4. ábra A Google találati listájának kezdete a „trendelemzések bibliometriai módszerekkel” témában

➤ Egyetemek, kutatóintézetek előszeretettel dolgoznak ki tudományometriai mutatószámokon alapuló értékelési sémákat. Az 5. ábrán bemutatott példa forrását szándékosan nem tüntetem fel, hogy megalkotói, felhasználói nehogy munkájuk bírálatának, „kipécézésének” tekintsék. Valójában százával található efféle értékelési „algoritmusok”, és nagy részük alighanem jó szolgálatot tesz a szóban forgó intézmény döntéshozói számára.

#### 3.2.4. Pályázati indikátor (IPa)

**IPa = Pa<sub>ö</sub> × M<sub>ö</sub> + Pa<sub>ny</sub> × M<sub>ny</sub>**, ahol **Pa<sub>ö</sub>** az összes beadott pályázat számának éves átlaga a PhD megszerzése óta (PhD hiányában az első publikáció megjelenési évétől); **Pa<sub>ny</sub>** a nyertes pályázatok számának éves átlaga a PhD megszerzése óta (PhD hiányában az első publikáció megjelenési évétől); **M<sub>ö</sub>** a beadott pályázatokra vonatkozó munkaköri szorzó; **M<sub>ny</sub>** a nyertes pályázatokra vonatkozó munkaköri szorzó

5. ábra Példa egy intézmény értékelési rendszerében használt mutatószámra

Két dolgot semmiképpen nem szabad azonban szem elől téveszteni. Az ilyen helyi használatra szánt mutatószámok akkor és azért hasznosak, ha a lehető legnagyobb mértékben figyelembe veszik a helyi adottságokat és igényeket. Ennek általában az az ára, hogy más közegben gyakorlatilag alig, vagy egyáltalán nem használhatók. Ezért ezek szélesebb körben való publikálásuk jobb esetben felesleges, rosszabb esetben félrevezető és káros. A másik tudnivaló az, ami néhány bekezdéssel korábban már szerepelt, de nem lehet elégszer hangsúlyozni: a legkiválóbban megalkotott mutatószámok is legfeljebb csak támogathatják a felelős emberi döntést, de semmiképpen sem helyettesíthetik.

➤ Egy másik értékelési célú tudományometriai konstrukciót nem hagyhatok inkognitóban. A Semmelweis Egyetem, Általános Orvostudományi Kar Bioinformatika Tanszéke által létrehozott tudomanymetria.com honlap (a

naprakész változat jelenlegi elérhetősége: [scientometrics.org/](http://scientometrics.org/)) a magyar kutatók tudományometriai mutatószámaihoz kísérel meg viszonyítási alapokat („benchmark”) megadni a Magyar Tudományos Akadémia osztályainak szakterületi felbontása szerint. A javasolt rendszer erős kritikákat kapott a tudománymetria művelőit és felhasználóit képviselő szerzőktől [Kamarás et al., 2021]. A szerzők véleménye szerint „a rendszer leegyszerűsítő, sematikus, és nem tudja a való élet problémáit kezelni [...]. Egyre több paraméter bevezetésével és a tudományterületek szerinti differenciálással éppen abba az irányba halad, hogy egységes 'pontrendszer' szerint nem kezelhető a tudományos teljesítmény. Innen már csak egy lépés eljutni oda, hogy az egyes tudományterületek sajátos szempontjait érvényesítő szakértői testületek tudják a pályázatokat a legjobban elbírálni – amelyben a számukra értő módon felkínált tudománymetria természetesen hasznos segítséget nyújthat.” A zárómondat egybevág az imént többször is deklarált alapelvvel.

#### ***1.4. A tudománymetria, mint üzlet***

Az utóbbi évtizedekben a könyvtári tevékenység súlypontja egyre inkább az információs szolgáltatások irányába tolódik el. A csökkenő költségvetési támogatások pótlására a könyvtáraknak ezeket a szolgáltatásokat mindinkább üzleti formában kell működtetniük. A tudományometriai/bibliometriai elemzések ilyen üzleti szolgáltatások tárgyai lehetnek. A Jülich Kutatóközpont Központi Könyvtára például az évezred elejétől próbálkozik ilyen szolgáltatásokkal [Ball & Tunger, 2006].

A könyvtáraknál sokkal nagyobb mértékben növekszik a kifejezetten erre a területre szakosodott vállalkozások által üzleti alapon készített tudományometriai/bibliometriai elemzések száma. Néhány fontos szereplő:

➤ Megtévészto neve ellenére a Eugene Garfield által alapított philadelphiai Institute for Scientific Information, a jelenleg a Clarivate által működtetett

Science Citation Index/Web of Science elindítója kezdetektől fogva üzleti alapon működött. Nemcsak a bibliográfiai adatbázisok, hanem a tudományometriai mutatószámokat tartalmazó csomagok (InCites, Journal Citation Report, Essential Science Indicators) külön-külön előfizetési díj, mégpedig igen magas előfizetési díj ellenében használhatók. Ezzel szemben az Elsevier tulajdonában lévő Scopus, bár bibliográfiai adatbázisának előfizetési díja a Web of Science-ével összemérhető, a tudományometriai mutatószámokat tartalmazó SCImago adatait ([www.scimagojr.com/](http://www.scimagojr.com/)), valamint a saját fejlesztésű folyóirat-indikátorait ([www.scopus.com/sources](http://www.scopus.com/sources)) mindenki számára ingyenesen hozzáférhetővé teszi.

➤ A tudománymetria egyik legelső európai kutatóközpontja, a Leideni Egyetem „Centre for Science and Technology Studies” egysége a bázisa a CWTS B.V. tanácsadó cégnek. A cég egyes kiadványai és termékei ingyenesen elérhetőek (pl. egy hasznos tanulmány az intézmény vezető kutatóitól: [Waltman & Noyons, 2018]; vagy az intézmény által kifejlesztett egyetemi világrangsor: [www.leidenranking.com/](http://www.leidenranking.com/)). A fő tevékenységük azonban a megrendelésre készített elemzések értékesítése.

➤ A kanadai alapítású Science-Metrix „független kutatásértékelési céggként” hirdeti magát, és azzal büszkélkedik, hogy 2002-es alapítása óta több mint 500 sikeres szerződéses elemzést készített a világ több tucat kutatási-fejlesztési intézményének. Működésének kezdeti szakaszában élen járt a tudománymetria társadalom- és humán tudományi alkalmazásának vizsgálatában is [Archambault & Vignola Gagné, 2004].

A cég „függetlenségét” némiképpen árnyalja az a tény, hogy 2018-ban az Elsevier megvásárolta.

Érdeemes megjegyezni, hogy a legutóbbi évtizedben a tudománymetria „modernebb” (főként az internetes jelenlétén alapuló) módszereire

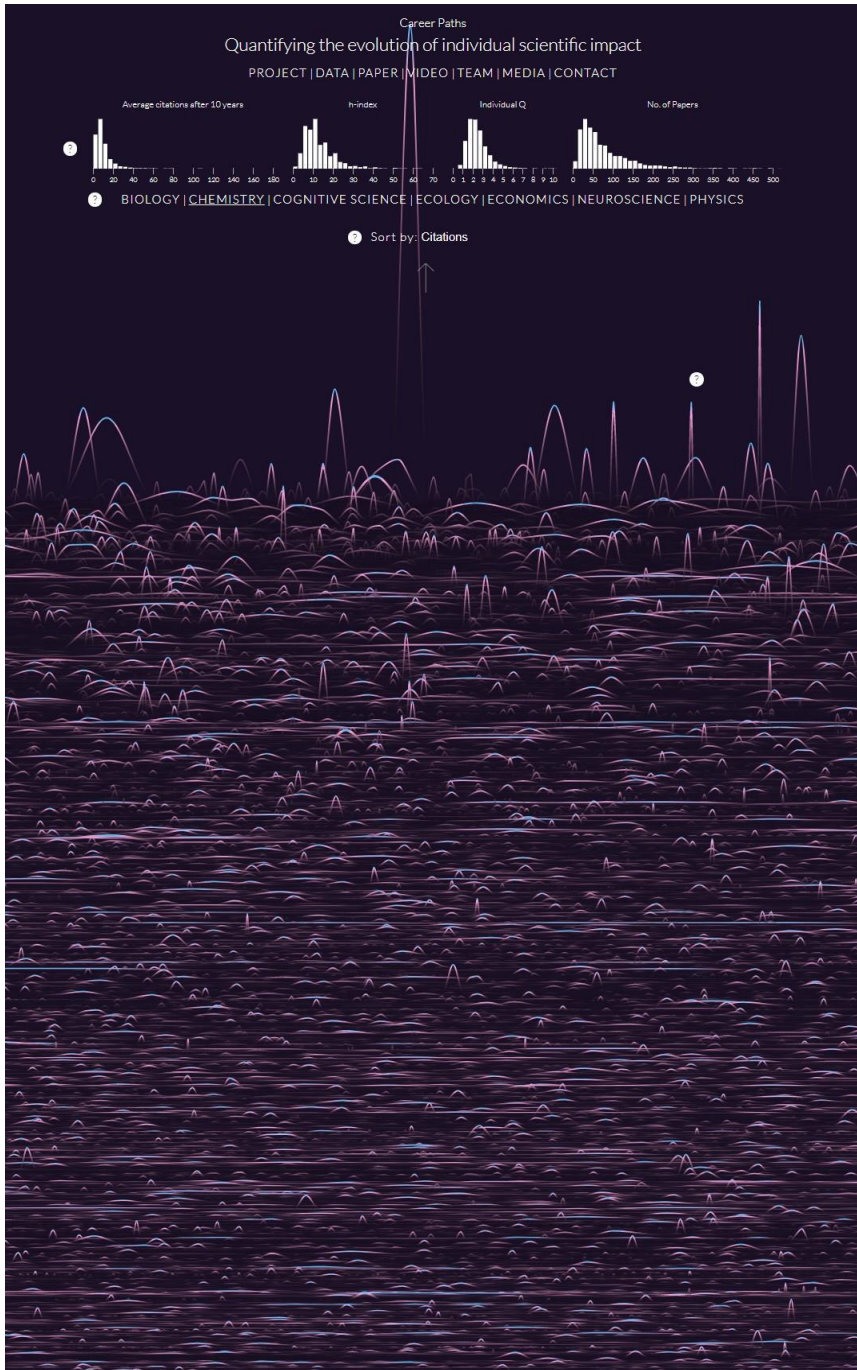
specializálódott vállalkozásai is megjelentek. A 2011-ben alapított Altmetric és a 2012-ben alapított Plum Analytics a tudományos kutatás eredményeinek a közösségi médián keresztül kifejtett hatását igyekeznek számszerűsíteni és ebből következtetéseket levonni. A vállalkozások tudományos megalapozottsága és eredményessége vitatható és vitatott, alapítóik üzleti sikere azonban vitathatatlan. Az Altmetric 2012-ben a Digital Science, a Plum Analytics 2017-ben az Elsevier portfóliójába került – nyilván nem elhanyagolható összegekért.

### KITÉRŐ

Lehet, hogy az olvasók közül van, aki úgy véli, hogy túlzott jelentőséget tulajdonítok a szóban forgó szolgáltatások tulajdonosi hátterének. Semmiképpen nem gondolom, hogy ennek a szakmai szempontból közömbösnek látszó tényezőnek meghatározó szerepe lenne a szolgáltatások működésében. De azt se felejtjük el, hogy a publikációs ipar legjelentősebb képviselői – az Elsevier, a Springer–Nature, a Wiley – „mindössze” egy nagyságrenddel képviselnek kisebb piaci értéket, mint a Big Pharma cégei, a nagy gyógyszeripari vállalatok. Az információs ipar óriásai, a Microsoft, a Google, a Facebook (illetve most már Meta) ugyanakkor a gyógyszercégek többszörösét érik. Ahogy a gyógyszeripar esetében is felvetődik a kérdés, hogy elsősorban a felhasználók érdekeit vagy saját üzleti érdeküket szolgálják, a kérdés nyitott az információs szolgáltatások, ezen belül a tudományos információt érintő szolgáltatások területén is. Ideális esetben természetesen ez a két érdek egybeeshet, de sajnos a világon egyre kevesebb tere van az idealizmusnak.

### **1.5. A tudománymetria, mint művészet**

Az elsőre talán meghökkentő társítás nem egészen légből kapott. A strukturális tudománymetria által vizsgált hálózatok, mintázatok vizualizációja esztétikailag is izgalmas képeket hozhat létre. A Barabási Albert László által vezetett bostoni Barabásilab „Rejtett mintázatok – A hálózati gondolkodás nyelve” című kiállítása jelenleg turnézik Európában (2020. október 10.–2021. március 21. között a budapesti Ludwig Múzeumban volt látható) [Barabásilab: Hidden patterns – The language of network thinking. Ludwig Museum, 10. October, 2020 – 21. March, 2021]. A teljes joggal művészi alkotásoknak tekinthető vizualizációk egy része tudománymetriai adatok (társszerzőség, idézetek) alapján készült. A 6. ábrán ábrán látható kép (az egyéni tudományos hatás változásának mennyiségi jellemzése) nem szerepel ugyan a kiállításon, de biztosan helyet kaphatna bármelyik modern galéria falán.



6. ábra Az egyéni tudományos hatás fejlődésének mennyiségi alakulása (a Barabásilab honlapjáról: [scienceofsuccess.barabasilab.com](http://scienceofsuccess.barabasilab.com))



### **1.6. A tudománymetria, mint sport**

Bár régi vesszőparipám a tudomány és a sport közötti hasonlóságok feltárása, semmilyen érvem nincs arra, hogy a tudományos kutatás bármelyik területét – akár magát a tudománymetriát – sportnak lehessen tekinteni. A sport és a tudománymetria közötti kapcsolat mindazonáltal sokrétű. A sporttal kapcsolatos tudománymetriai kutatásokról többszáz cikket találunk az irodalomban. A sport területére kifejlesztett mutatószámokat (pl. a sakk Élő pontszámát) fel lehet használni a tudománymetriában [Franceschet & Colavizza, 2017; Lehmann & Wohlrabe, 2017], és a fordítottjára is lehet példát találni: a Hirsch-féle h-index jól használható a krikett ütőjátékosok eredményességének értékelésében [Prathap, 2009; Daud et al., 2015].

### **1.7. És még??**

További ötleteknek csak az emberi fantázia szab határt. Főként a kínaiból angolra fordított irodalomban fordul elő, hogy a tudománymetriát „scientometrics” vagy esetleg „scientometry” helyett „scientology”-nak fordítják – így akár vallás is lehet belőle. Természetesen lehet hobbi vagy akár szenvedély is – magam is ismerek olyanokat, akiket aligha lehet belőle kigyógyítani.

Végül pedig – mint akármi más – a tudománymetria humorforrás is lehet (lásd 7. ábra).



7. ábra „A kristálygömb hókusz-pókusszal parancsolja, vagy tudománymetriai alapon?”

#### KITÉRŐ

Ezt a (Sidney Harris rajza nyomán készült) karikatúrát évtizedeken keresztül használtam tudományos és ismeretterjesztő előadásaimban. Egy kínai utazásom során döbrentem rá, hogy egy ilyen „ártatlan tréfa” is mennyire kultúrafüggő. A kínai hallgatóság számára a stilizált számítógép még csak-csak felismerhető volt, de a jósnővel és a kristálygömbbel külön hosszadalmas magyarázat nélkül nem tudtak mit kezdeni.

## 2. A tudománymetria története

### 2.1. A tudománymetria úttörői

Amikor egy kutatási terület intézményesülni kezd, az első lépések egyikeként megteremti a saját mítoszát. Olyan "úttörőket" talál, akiknek tiszteletre méltó képét büszkén teheti közszemlére; a dicső múlt révén megelőlegezve a még dicsőbb jövőt. Talán nem véletlen, hogy ezeket az ikonikus figurákat általában holtuk után éri ez a megtiszteltetés, amikor már semmiképpen nincs módjuk tiltakozni ellene.

Amikor a tudománymetria az 1960-as években elindult ezen az úton, nem cselekedett másként. A legvakmerőbbek a családfát egészen a XIX. század második feléig, Alphonse de Candolle-ig [Szabó, 1985], vagy akár az első feléig, Friedrich Engelsig [Schubert, 2019] vezették vissza.

#### KITÉRŐ

Alphonse Pyramus de Candolle (1806–1893) francia-svájci botanikus 1873-ban megjelent *Histoire des sciences et des savants* (A tudomány és a tudósok története) című könyvében kiterjedten használt statisztikai táblázatokat a tudomány történetének, a tudomány fejlődését befolyásoló különféle tényezőknek a vizsgálatához. Provokatív megállapításokat is tett, például a nemzetközi tudós társaságok, akadémiák tagságának vizsgálatából azt a következtetést vonta le, hogy a protestánsok nagyobb tudományos aktivitást fejtettek ki, mint a katolikusok.

Ami ezeknél a történeti kalandozásoknál lényegesebb, a tudománymetria hamar megtalálta azokat a koherensnek látszó alapokat, amelyekre építkezhet: három matematikailag megfogalmazható összefüggést: Lotka, Bradford és Zipf törvényeit.

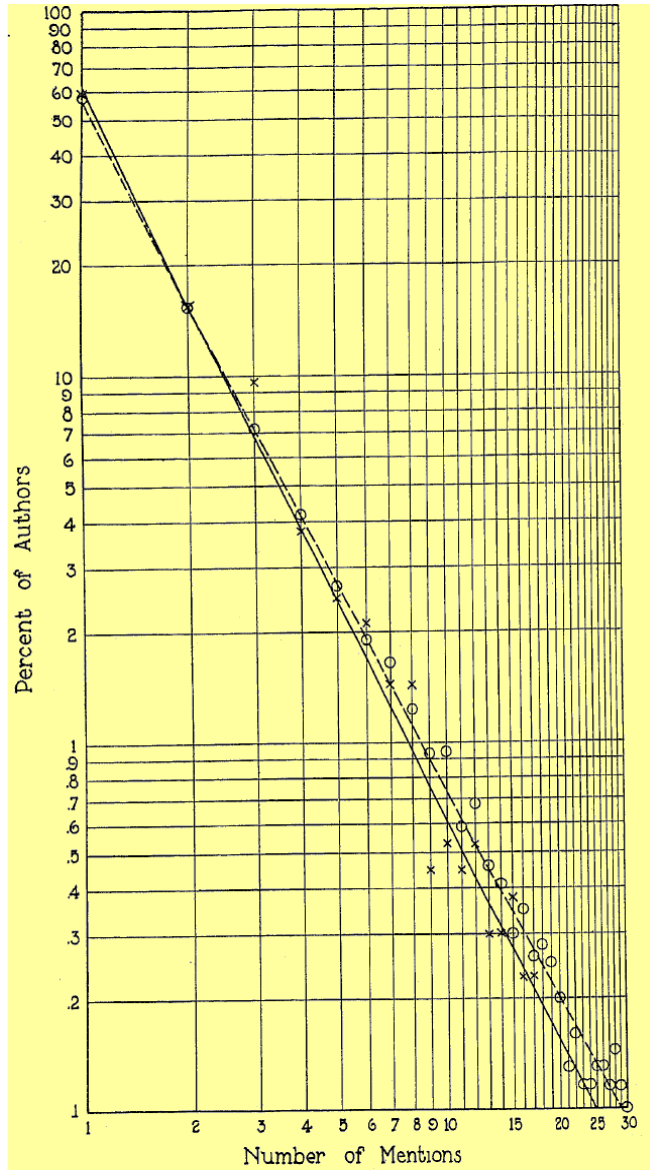
Függetlenül e törvények érvényességétől, vagy a tudománymetria kialakulásában és fejlődésében betöltött tényleges szerepüktől, érdemes megismerkedni e három törvénnyel és névadóikkal, mert mind a hárman – bár egymástól igen különböző módon – kivételes személyiségek voltak [Schubert & Schubert, 2016].

➤ Alfred J. Lotka (1880–1949) (8. ábra) amerikai szülők gyermeke volt. Tanulmányait Párizsban kezdte, Birminghamben folytatta, majd Lipcsében tanult kémiát. Az Egyesült Államokba visszatérve fizikát hallgatott és oktatott, vegyészként, majd biológus kutatóként dolgozott. 1924 től nyugdíjazásáig a Metropolitan Life Insurance Company statisztikusa volt.



8 ábra Alfred James Lotka (forrás: Wikipedia)

Úttörő cikkében [Lotka, 1926] névmutatókból választott mintákon vizsgálta meg a szerzői produktivitás gyakoriságeloszlását, vagyis, hogy hány szerző írt 1, 2, 3, stb. cikket. Tapasztalata szerint az  $x$  cikket írt szerzők száma  $x^{-n}$  nel arányos, ahol  $n$  egy, az eloszlásra jellemző állandó, amelyet jó közelítéssel 2-nek talált (9. ábra). A jelenség magyarázatával a cikk nem foglalkozott. A gondolatnak Lotka munkásságában sem előzménye, sem következménye nem volt, de mások gondolkodását alaposan megmozgatta. 1945 óta a Web of Science szerint közel 1000 cikkben idézték az 1926-os munkát.



9. ábra Lotka törvénye (forrás: [Lotka, 1926])

Tevékenysége nem korlátozódott a tudományometriára. Nevéhez fűződik a kétkomponensű kémiai oszcillációk Lotka–Volterra modellje; *Elements of Physical Biology* című könyve [Lotka, 1925] az 1940-es években a matematikai biológia alapvetésévé vált; őt tekintik továbbá a statisztikai demográfia atyjának is.

## KITÉRŐ

A Lotka-Volterra modell messze túlnőtt a kémiai kinetika területén. A két komponens lehet két állatfaj, két gazdasági szereplő vagy éppen két tudományos eszme. A modellben egyik a „ragadozó”, a másik a „zsákmány” szerepét játssza. Adott feltételek mellett a két komponens mennyisége felváltva növekszik és csökken, vagyis oszcillál. A modellről az eltelt évtizedekben többtízezer cikket írtak. Alfred Lotkának ez a még szinte diákkorában támadt ötlete kétségtelenül egyik legnagyobb hatású tudományos eredménye lett.

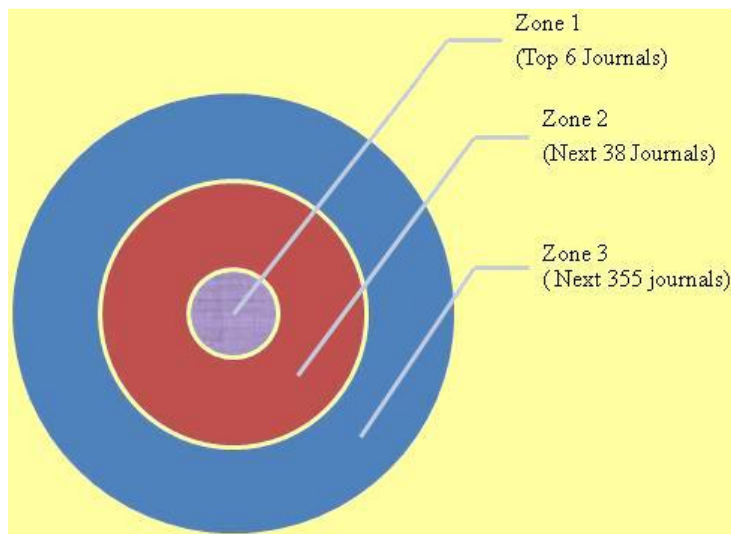
Lotka és Volterra semmiféle kapcsolatban nem álltak egymással, csak a modell neve kapcsolja őket össze. Vito Volterra (1860–1940), olasz matematikus, biológiai jelenségek leírásához alkotott Lotkáéhoz hasonló modellt. Volterra, zsidó származású lévén, heves ellenzője volt Mussolini fasiszta rendszerének. Azt vallotta: „Birodalmak emelkednek és buknak, de Eukleidész tételei örökké ifjak maradnak.”

➤ Samuel C. Bradford (1878–1948) (10. ábra) esti tagozaton, vegyészként szerezte PhD-jét, de 21 éves korától nyugdíjazásáig a Londoni Tudományos Múzeum könyvtárosaként dolgozott. Már mindig elegáns öltözködése, megjelenése is tiszteletet parancsolt. Gomblyukából sohasem hiányzott egy szál rózsaszín.



10. ábra Samuel Clement Bradford (a Google Scholar fotója nyomán)

Klasszikussá vált cikkében azt vizsgálta, hogy hogyan oszlanak meg egy szakterület legfontosabb cikkei (tételek) a folyóiratok (források) között. Azt találta, hogy ha a forrásokat a tételek száma szerint sorba rendezzük, akkor a legtöbb tételt tartalmazó, a középső és a legkevesebb tételt tartalmazó harmad elemeinek száma  $1 : n : n^2$  arányban áll egymással, ahol  $n$  az úgynevezett Bradford állandó (11. ábra). Az összefüggést Bradford gyakorlati célra szánta, semmiféle magyarázatát nem adta, és azóta sem találták.



11. ábra Bradford törvénye

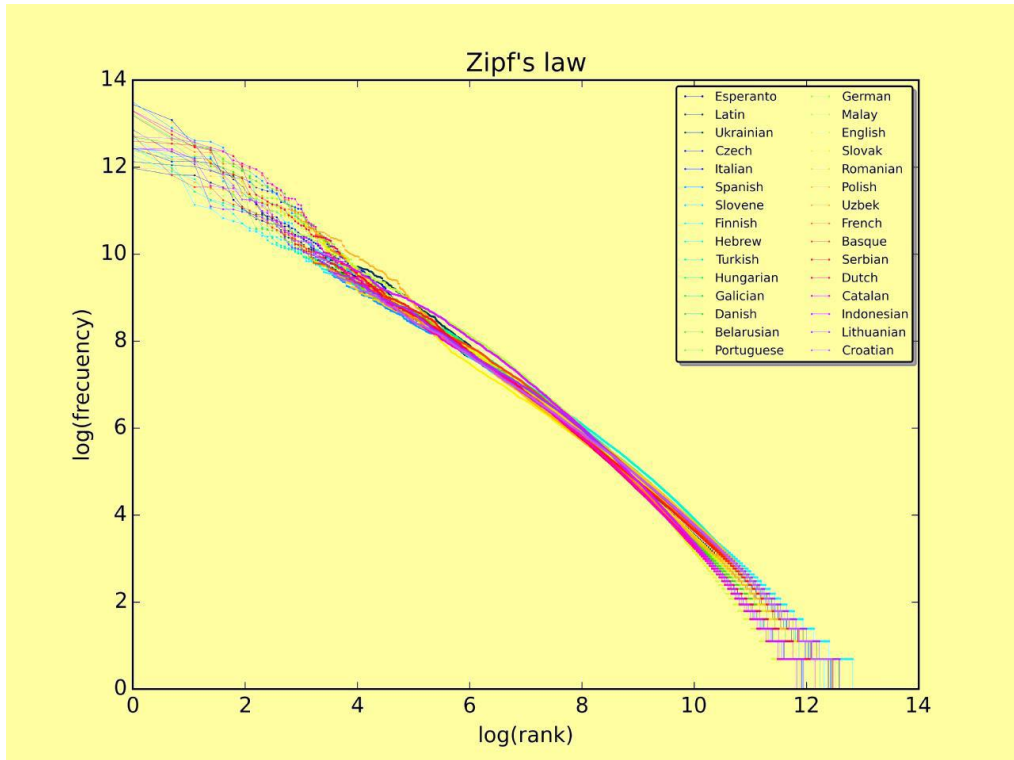
Nyugdíjba vonulása utáni éveit Bradford kedvteléseinek: a rózsák nevelésének és a tudománynak szentelte. Ekkor jelent meg két könyve: *A rózsák románca* [Bradford, 1946] és a *Dokumentáció* [Bradford, 1948]. Az előbbit a botanika, utóbbit az információtudomány gyűjteményei tartják számon.

➤ George K. Zipf (1902–1950) (12. ábra) a Harvardon tanult (rövid berlini és bonni kitérővel), és ott is dolgozott haláláig. Alapterülete a nyelvészet volt, itt megfogalmazott gondolatait kísérte meg világmagyarázó elméletté általánosítani.



12. ábra George Kingsley Zipf (Fabian Bachrach fotója; *Glottometrics*, 2002)  
1935-ben használja először azt az összefüggést, amelyet ma a tudománymetria Zipf-törvényként tisztel. Eszerint egy csökkenő gyakorisági sorba rendezett szóhalmazban a rangsorban  $x$ -edik helyen álló szó előfordulási gyakorisága  $x^{-n}$ -nel arányos, ahol  $n$  egy, az eloszlásra jellemző állandó, amelyet Zipf jó közelítéssel 1-nek talált [Zipf, 1935]. Az összefüggést azóta számtalan gyakorlati példán közelítőleg igazolták (lásd pl. 13. ábra).





13. ábra Zipf törvénye a Wikipédiák szógyakoriság eloszlására

A kor nyelvészei munkáit szakmailag elfogadhatatlannak minősítették. Politikai nézeteiért is támadták: német felmenőkkel rendelkező családban született; ennek bizonyára szerepe volt a német nyelv iránti érdeklődésében és Hitler eszméi iránti kendőzetlen rajongásában. Zipfet nem tántorította vissza a szakma értetlensége, szívósan dolgozott általános elméletén, és 1949-ben megjelent főműve: *Az emberi viselkedés és a legkisebb erő kifejtés elve* [Zipf, 1949]. Az utókor megbocsátóbb. Az ideológiai tévelygések elhomályosulnak, a kortársak értetlensége a meg nem értett zsenik sajátja. Az örült ötletek értő kezekben meggyőző eredményeket ihlettek. Herbert Simon, Benoit Mandelbrot, Barabási László új értelmet adott Zipf zavaros elgondolásainak. 1945 óta a *Web of Science* szerint műveit kb. 3000 cikkben kb. 6000 alkalommal idézték.

## KITÉRŐ

Guglielmo Ferrero (1870–1942) olasz liberális filozófus, történetíró és író 1894-ben megjelent cikkének címe: „A szellemi tehetetlenség és a legkisebb erő kifejtés törvénye”. Sokan ezt a cikket tekintik Zipf munkássága szellemi elődjének, bár a valóságban kevés közös gondolatot lehet találni a két műben, és Zipf semmilyen módon nem is hivatkozik Ferreróra. Ferrero végkövetkeztetésével viszont alighanem sokan egyetértenek: „Az ember irtózása a fizikai és szellemi erő kifejtéstől könnyen megmagyarázható. Minden munka a test szöveteinek bomlásával jár, és ez fájdalmat okoz. Minél gyengébb a szövet, annál gyorsabb a bomlás. A legtöbb ember agya elég gyenge ahhoz, hogy hamar kimerüljön a munkában.” Azért világmagyarázatnak ez egy kicsit kevés.

A legkisebb erő kifejtés elve napjainkban az információ- és könyvtártudományban talált otthon. A gyors és könnyű információkeresés biztosítása a korszerű könyvtári és információs szolgáltatások egyik alapkövetelménye; ehhez nyújthat segítséget az elv értő alkalmazása.

### **2.2. A tudománymetria névadói**

Az előfutárok vagy úttörők után a névadókról is megemlékezhetünk.

1969-ben, szinte egyidőben jelent meg az angol Alan Pritchard cikke [Pritchard, 1969], és a szovjet Vaszilij V. Nalimov és Zinajda M. Mulcsenko könyve [Nalimov & Mulcsenko, 1969]. Ezt a két művet szokás a *bibliometria*, illetve a *tudománymetria* kifejezések forrásának tekinteni.

Érdeemes megjegyezni, hogy a belga polihisztor, Paul Marie Ghislain Otlet 1934-ben megjelent könyvében [Otlet, 1934] már használta a francia „bibliometrie” szót, mégpedig a máig is használatos értelemben.

## KITÉRŐ

Paul Marie Ghislain Otlet (1868–1944) belga író, vállalkozó, ügyvéd és békeharcos volt. Sokan az információtudomány egyik megalkotójának és a dokumentáció szakterülete atyjának tekintik. Dewey rendszerének továbbfejlesztésével Henri La Fontaine-nel együtt alkották meg az Egyetemes Tizedes Osztályozást (ETO).

### **2.3. A tudománymetria alapító atyái**

Az előfutárok, úttörők, névadók fantáziája nem lett volna elég ahhoz, hogy ötleteikből egy tudományterület jöjjön létre. Ehhez szükség volt olyan emberekre, akik megtervezték és felépítették azt a szellemi konstrukciót, amely a tudománymetria valóságos létrejöttét lehetővé tette. A tudománymetria alapító atyáinak Derek John de Solla Price-t és Eugene Garfieldet szokás tekinteni; én mindenképpen hozzátennék még egy nevet: Braun Tiborét.

➤ Derek John de Solla Price (1922–1983) (14. ábra) amerikai fizikus, tudománytörténész *Kis tudomány, nagy tudomány* című könyvét [Price, 1963] tekintik a mai értelemben vett tudománymetria alapvetésének.



14. ábra Derek John de Solla Price (forrás: Wikimedia Commons)

Emlékét sok minden egyéb mellett a *Scientometrics* folyóirat által alapított emlékérem (Derek de Solla Price Memorial Award) őrzi (15. ábra). A díjat kétévenként ítélik oda a tudomány kvantitatív vizsgálatában legkiemelkedőbb eredményeket elért kutatóknak.



15. ábra A Derek de Solla Price emlékérem  
(Szöllősy Enikő alkotása)

➤ Eugene Garfield (1925–2017) (16. ábra) amerikai vegyész, nyelvész, információkutató 1955-ben alapította meg cégét, az Institute for Scientific Informationt (mai tulajdonosa a Clarivate), melynek Science Citation Index nevű terméke az idézettségvizsgálatok lehetővé tételével a tudományometriai kutatásoknak máig is egyik legfontosabb adatbázisa.



16. ábra Eugene Garfield (forrás: Wikimedia Commons)

## KITÉRŐ

A vegyész végzettségű Garfield az információtudománnyal való első találkozásai során ismerkedett meg az amerikai jogszolgáltatás egyik alapvető forrásanyagával, a Shepard's Citations című kiadvánnyal. Ezekben az 1873 óta kiadott kötetekben pontosan követhető, hogyan hivatkoznak az Egyesült Államok bírósági ítéletei a korábbi hasonló ügyek peranyagaira. (Az Egyesült Államokban precedensjog érvényesül, vagyis a korábbi hasonló esetekben született döntések irányadóak a későbbi bírósági ítéletekre.) Garfield felismerte, hogy a jogi információk visszakövetésének ez a módszere, amit az amerikai jogi szakzsargon „shepardizálásnak” is nevez, átvihető a tudományos információk forrásainak visszakövetésére. Ehhez a tudományos cikkek hivatkozásainak olyan indexelésére lenne szükség, ahogy azt a Shepard's Citations teszi a jogi előzményekkel. 1955-ben publikálta a Science Citation Index (SCI) koncepcióját [Garfield, 1955], és 1957-től hozzájárult annak gyakorlati megvalósításához. 1964-ben jelent meg az első SCI közel másfélmillió hivatkozás feldolgozásával öt nyomtatott kötetben. Az adatbázis jelenlegi változata, a Web of Science, kétmilliárd indexelt hivatkozást tartalmaz.

➤ Braun Tibor (1932–) (17. ábra) magyar vegyész, az ELTE professzora, információkutató, folyóiratszerkesztő a *Scientometrics* című folyóirat alapító főszerkesztőjeként nagy lépéssel járult hozzá a tudománymetria intézményesüléséhez, és az MTA Könyvtárának igazgatóhelyetteseként a magyar tudománymetriai műhely létrehozásához.



17. ábra Braun Tibor (forrás: személyes archívum)

### KITÉRŐ

A 17. ábra fotóján Braun professzor kezében nem futball-labda van. A szabályos öt- és hatszögekből álló csonkított ikozaéder valójában egy fullerén molekula modellje. Ez a molekula Braun Tibor kémiai kutatásainak egyik témája. Ő alapította 1993-ban a *Fullerenes* című máig is sikeres nemzetközi folyóiratot (jelenlegi címe: *Fullerenes, Nanotubes and Carbon Nanostructures*). Ez a furcsa alakú és furcsa nevű, hatvan szénatomból álló molekula R. Buckminster Fuller (1895–1983) nevét őrzi. Fuller építész, tervező, feltaláló, író, költő és futurológus volt. Az építészeti „védjegyének” számító geodetikus kupolák ikozaéderes szerkezete sugallta a különleges szénmolekula elnevezését.

## **2.4. A tudománymetria Magyarországon**

A tudománymetria magyarországi elfogadtatásának mérföldköve volt két könyv, Derek de Solla Price és Vaszilij V. Nalimov munkáinak lefordíttatása és megjelentetése.

Price könyvének akkor már jelentős visszhangja volt a „nyugati világban”, de magyar kiadása úttörő tetteknek számított a hivatalos tudományfilozófiával korántsem egyező szemlélete miatt. Ezen felül a magyar kiadás szerkesztésében részt vevő Braun Tibor és Ruff Imre a kötetet kiegészítette Price több, másutt megjelent cikkével, amelyek a könyv egy későbbi amerikai kiadásában is helyet kaptak [Price, 1985].

Nalimov és Mulcsenko könyve voltaképpen a tudománymetria szó ősforrása. A „scientometrics” a „naukometriya” fordításaként jelent meg először az angol nyelvű szakirodalomban. Rám a könyv azért is volt nagy hatással, mert míg korábban magam is vallottam a Disraelinek tulajdonított „kis hazugság – nagy hazugság – statisztika” közhelyet, ebből a könyvből értettem meg, hogy sokszor csakis a statisztika eszközével lehet az igazságot kimondani. Nalimovék statisztikai olyan meggyőző erővel tárják fel a kor szovjet tudományának zártságát és belterjességét, és ennek negatív hatásait, ami más módon szakmailag és politikailag sem lett volna lehetséges [Schubert, 2014]. (Nalimov, aki korábban éveket töltött a Gulágon, bizonyára mérlegelte a lehetőségeit.)

A két könyv természetesen „csak” népszerűsítő eszköz volt. A magyarországi tudománymetria „húsát és véré” a Science Citation Index és az annak alapján végzett elemző munka adta, amelynek magam is részese lehettem. Braun Tibor elérte, hogy az MTA Könyvtára információs szolgáltatás (ASCA) céljaira megvásárolja az adatbázis több éves anyagának nyomtatott kötetait és mágnesszalagjait, és a hetenkénti frissítéseket. A sikeres (és a könyvtár

számára jól jövedelmező) szolgáltatás „melléktermékeként” fogadta be a könyvtár a tudományometriai tevékenységet. Az SCI adatokkal a tudományometriai elemzésekhez szükséges adatoknak egy olyan kincsesbányájához jutottunk, amilyen a philadelphiai anyaintézményen kívül senki másnak nem állt rendelkezésére. Emellett megvolt az a szellemi és számítástechnikai kapacitás is, amivel ezekből az adatokból érdekes és értékes eredményeket tudtunk kiszűrni. Akkor még nem volt ismeretes az „adatbányászat” kifejezés, de voltaképpen azt csináltuk. Az MTA SzTAKI IBM 3031 számítógépe a keleti blokk országait sújtó embargó enyhítése nyomán Kelet-Európa legnagyobb számítógépeként került Magyarországra. Ezen a gépen kaptunk (főleg éjszaka) gépidőt munkánk végzésére. A mai szemmel nézve nevetséges mennyiségű (néhány száz megabájtnyi) adat sok tucat nagy mágnesszalagról került a hordónyi méretű merevlemezekre, úgy futtathattuk az előzetesen lyukkártyán leadott programjainkat. Programozási vagy adathiba esetén másnap éjszaka kezdhettük az egészet előlről. Fáradságos munka volt, de még mindig nagyságrendekkel hatékonyabb, mint ameddig a nyomtatott kötetekből kézzel kellett kikeresnünk és számlálnunk az adatokat. 1983-ban jelent meg az első átfogó elemzés, amit az SCI adatok számítógépes feldolgozásával készítettünk.

Az 1980-as évektől napjainkig az MTA Könyvtárában működő tudományometriai „műhely” munkatársai 300 körüli tudományos cikket publikáltak, melyek összes idézettsége meghaladja a 10000-et.

A Derek de Solla Price emlékérmét eddig négy magyar kutató nyerte el: Braun Tibor 1986-ban, Schubert András 1993-ban, Glänzel Wolfgang 1999-ben, Vinkler Péter pedig 2009-ben.



## **3. A tudományometriai mutatószámok elemei**

Amint láttuk, a tudománymetria definíciója – a tudományos kutatás kvantitatív vizsgálata – igen tág, a leggyakrabban azonban a tudománymetria (bibliometriai) vizsgálatok tárgyát a tudományos publikációk, illetve azok metaadatai alkotják. Bár ezek az adatok (a cikk címe, szerzői, a megjelenés ideje stb.) látszólag nyilvánvalóak, a tudományometriai elemzésekben való használatuk sok problémát vet fel.

### ***3. 1. Szerzők, társszerzők***

A tudományos publikációkban általában egyértelműen feltüntetik a szerző(k) nevét. Neve feltüntetésével a szerző felelősséget vállal a cikk tartalmáért, egyszersmind dokumentálja (és ezzel „mérhetővé” teszi) a kutatásban való részvételét. Az álnéven (szerzői néven) való publikálás a tudományokban ma már ritka, de előfordulhat. A szerző személyének azonosítása azonban gyakran nem könnyű. A névazonosságok (homonímia) gyakoriak, egyes nyelvekben (japán, koreai) különösen. A nevek egyértelműsítése ma már a könyvtár- és információtudomány fontos területe, évente több ezer cikk jelenik meg ebben a témában, amelyben a mesterséges intelligencia és az adatbázisépítés (ORCID) egyaránt szerepet kap. A névazonosságokkal ellentétes probléma az, amikor egyazon szerző több különböző névváltozattal publikál. Ennek oka lehet a közepső névbetűk következtelen használata, névváltoztatás (házastársi nevek), a spanyol neveknél az anyai név következtelen használata, vagy a nem latin betűs nevek különféle átírása. Ezt a problémát biztonságosan csak a szerzői névadatbázisok (ORCID) használatával lehet megoldani.

## KITÉRŐ

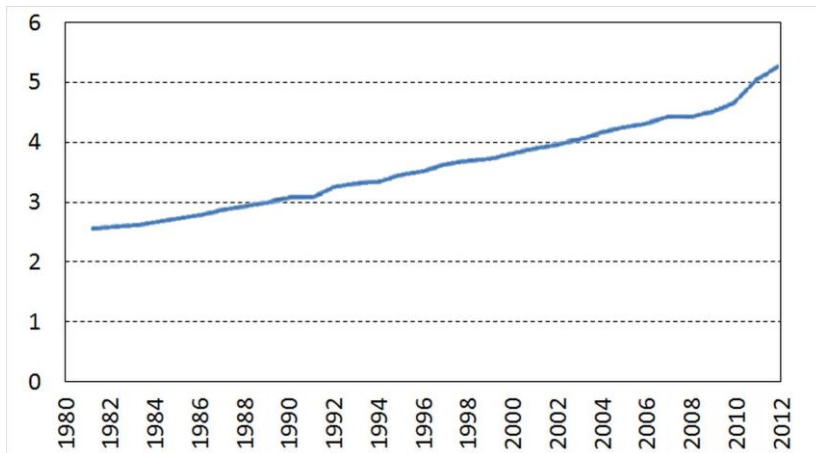
Az álnéven való publikálás a tudomány története során rendszeresen előfordult. Már Sir Isaac Newton is használt álnevet: ez Jehovah Sanctus Unus volt, amely nevének kissé módosított latinus Jsaacus Neuutonus változatának az anagrammája. A statisztika jól ismert t-próbájának megalkotója „Student” volt; e mögött az álnév mögött William Sealy Gosset angol statisztikus, vegyész és sörfőző rejtőzött. Elképzelhető, hogy a Guinness vezető sörmestereként rossz hírbe kerülhetett volna, ha fény derül titkos matematikusi szenvedélyére. A matematikus Sophie Germain szerzői álneve Monsieur Antoine Auguste Le Blanc volt. Ő alighanem női mivoltát akarta elrejtteni, hogy a kollégái komolyan vegyék munkáit. Ennek ellenpárjaként, a matematika és számítástudomány élő klasszikusa, Donald Knuth egy alkalommal Ursula N. Owens néven küldött be egy kéziratot. Utólag bevallotta, hogy szerette volna, hogy ezt az írását különleges alapossággal bírálják el.

Az utóbbi évtizedekben elszaporodó tudományos blogokban az írások jelentős része álnév (becenév, „nick”) alatt jelenik meg. Az álnevek mögött gyakran valódi nevükön sokat és eredményesen publikáló kutatók rejtőznek, akiknek kiléte legfeljebb csak egy szűk csoport számára ismert. Akadt olyan blogger, aki a becenevével jegyezte komoly tudományos folyóiratban megjelent néhány cikkét. Ez azonban a tudományos közösségben erős ellenérzést szült, és a jelenleg uralkodó nézet szerint ellentmond a COPE (Committee on Publication Ethics) alapelveinek.

A 2012-ben elindított ORCID (Open Researcher and Contributor ID; orcid.org) minden regisztrált szerzőt egy 16-karakteres egyedi azonosítóval lát el, ami egyben a profil weboldal URL-je (az enyém pl. orcid.org/0000-0001-8348-1775). A nonprofit kezdeményezéshez rövid idő alatt csatlakoztak a

legjelentősebb tudományos kiadók, egyetemek és kutatóintézmények, valamint természetesen maguk a kutatók is. Mára a regisztrált kutatók száma meghaladja a 10 milliót. Az azonosító használata nagymértékben megkönnyíti a személyekre vonatkozó statisztikák, elemzések készítését.

Jól ismert tény, hogy a tudományos cikkek szerzőinek száma egyre növekszik (18. ábra). A tudományos munka egyre inkább csapatmunkát igényel, és a résztvevők mindegyike igényt tart a nevének feltüntetésére. Ahogy említettük, ennek az oka nemcsak a hiúság (az is), hanem az, hogy dokumentálja a kutatómunkában való részvételét, és vállalja a publikált eredmények rá eső részéért a felelősséget is.



18. ábra A folyóiratcikkek átlagos szerzőszáma a Web of Science adatbázisban

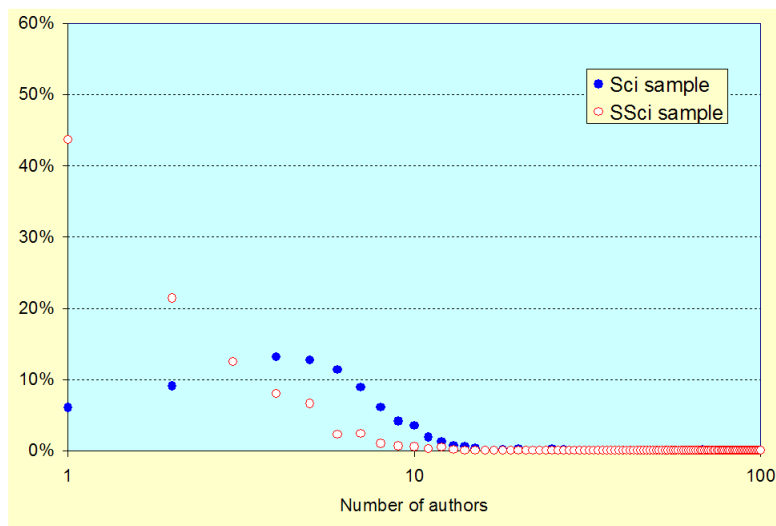
A társszerzőség mértéke szakterületenként erősen különbözik. Általában a „lágyabb” tudományterületeken (társadalom- és humán tudományok) és a matematikában alacsonyabb, mint a „keményebb” területeken (természet- és műszaki tudományok) (19. ábra).

Az utóbbi évtizedben megjelent a szerzőségnek egy új, szélsőséges változata: a „megaszerzőség”. Ilyenkor a cikk szerzőinek száma százas, esetenként ezres

nagyságrendű. Ez főként a részecskefizika és a molekuláris genetika területén fordul elő, ahol néha évekig, változó összetételű kutatógárdával készül el egy-egy publikáció. Ez a jelenség már olyan technikai problémákat is felvet, hogy az összes szerző nevét a folyóiratban nem tudják kinyomtatni, illetve a metaadatokat tároló adatbázisban a szerzők száma túllépi a rendelkezésre álló korlátot. A szerző neveként ilyenkor néha csak a szerzői kollektíva megnevezését adják meg, és külön dokumentumban közlik a kollektívában részt vevő kutatók névsorát.

### KITÉRŐ

A szerzői kollektíva nevén történő publikálást akár a szerzői álnév speciális eseteként is tekinthetnénk. Persze, nem az, mert ebben az esetben minden egyes szerző neve a cikk minden olvasója számára megtalálható a kiegészítő dokumentumban. Történeti példákat találunk azonban a „kollektív álnév” esetére is. A leghíresebb közülük Nicolas Bourbaki neve. Ezt a fiktív nevet fiatal francia matematikusok egy csoportja vette fel, akik az 1935-ös alakuló kongresszusukat követően 1939-től jelentették meg *Éléments de mathématique* (A matematika elemei) címmel a matematika mindenkori legkorszerűbb ismereteit tükröző értekezéseiket. A csoport változó résztvevőkkel mindmáig működik, a sorozat legutóbbi kötete 2016-ban jelent meg. A csoportban eddig közel ötven kutató működött közre.



19. ábra A cikkek számának szerzőszám szerinti megoszlása az SSCI és az SSCI adatbázisban [Schubert, 2016]

A tudományos teljesítmény értékelésekor fontos lenne tudni, hogy az egyes szerzők milyen mértékben járultak hozzá a társszerzős cikkben foglalt eredmények létrehozásához. Sajnos erre vonatkozó megbízható információk alig állnak rendelkezésre. Vinkler Péter kérdőíves vizsgálatot végzett vegyész kutatók között, hogy egy kiválasztott cikkmintában ki hány százalékra becsüli a saját hozzájárulását. Nem túl meglepő módon a becsült százalékok összege minden esetben messze meghaladta a 100%-ot.

Közelítő arányokat lehet megbecsülni a szerzők sorrendje alapján. A szerzők sorrendjére vonatkozó szokások, hagyományok szakterületenként különböznek. Az orvosbiológiai publikációk szerzői között általában a kísérleti munkát vezető kutató áll az első, az intézményt vezető kutató az utolsó helyen. A többi szerző neve többnyire „fontossági sorrendet” követ. Más területeken (pl. kémia) a szerzői sorrend nagyjából a hozzájárulási sorrendet követi. A matematikában és a fizikában a szerzők nevét legtöbbször névsor szerint tüntetik fel, így a sorrend semmilyen releváns információt nem tartalmaz.

A szerzői hozzájárulás kérdésének szélsőséges esete az, amikor kétségessé válik, hogy a szerzőként feltüntetett kutatók egyike-másika egyáltalán részt vett-e a kutatásban. Ez tudományetikai kérdés, aminek természetesen a tudománymetriai értékelések szempontjából is nagy jelentősége van. A „tiszteletbeli” szerzőséget befolyásos kutatók olykor kikövetelik maguknak, máskor éppen az érdemi szerzők kezdeményezik olyan "húzónevek" társszerzői részvételét, akiktől a cikk kedvezőbb elbírálását vagy fogadtatását remélik.

A tudományos kiadók, szerkesztőségek, bizottságok igyekeznek alaposan és pontosan megfogalmazni azokat a kritériumokat, amelyek a szerzőséghez szükségesek. Az ICMJE (International Committee of Medical Journal Editors) ajánlásait szerint szerzőnek az a személy tekinthető, aki

- 1) lényegesen hozzájárult a munka kiötléséhez vagy megtervezéséhez, illetve az adatok gyűjtéséhez, elemzéséhez vagy értékeléséhez; ÉS
- 2) részt vett a cikk megfogalmazásában vagy önálló szellemi tartalommal járult hozzá a végleges forma kialakításához; ÉS
- 3) jóváhagyta a publikálásra kész változatot; ÉS
- 4) hozzájárult ahhoz, hogy számonkérhető legyen a cikk bármely részének pontosságával vagy alaposságával kapcsolatos kérdések megfelelő vizsgálatának vagy megoldásának biztosításáért.

Egy másik újabb próbálkozás a CRediT (Contributor Roles Taxonomy), amelyet az azóta megszűnt kanadai CASRAI (Consortia Advancing Standards in Research Administration Information) kezdeményezett. Jelenleg a NISO (National Information Standards Organization) égisze alatt működik ([credit.niso.org](http://credit.niso.org)). Ez a taxonómia a publikációs hozzájárulás 14 alaptípusát különbözteti meg, és a javaslat szerint a cikkekben meg kellene adni minden egyes szerző hozzájárulásának a típusát. A CRediT honlapja szerint a kiadók hosszú sora csatlakozott a javaslathoz.

A kritériumok lehetnek akármilyen pontosak és szigorúak, ellenőrzésük, betartatásuk reménytelen. Az általános vélekedés szerint a nem kellően megalapozott érdemekkel rendelkező "fantomszerzők" aránya a tudományos cikkek szerzői között az utóbbi évtizedekben jelentősen nőtt, ha talán nem is az átlagos szerzőszám növekedéssel azonos mértékben.

Egy kérdés komolyságát mi sem jelzi jobban, mint hogy humor forrásává is válik. A tudományos humor kimeríthetetlen forrása, az *Annals of Improbable Research* ([www.improbable.com/](http://www.improbable.com/)) közölte Lundberg és "munkatársai" (az idézőjel jelentősége hamarosan kiderül) cikkét [Lundberg et al., 2016], amely azt sugallja, hogy a szerzők számának emelkedése az entrópia sajnálatos növekedését eredményezi. (Nem teszik hozzá, csak én következtetem, hogy ezzel talán a klímaváltozásért is felelősség terheli...) Javaslatként – a folyóirat alapelvehez hűen tréfásan, de elgondolkodtatva – háromra korlátoznák egy cikk lehetséges szerzőinek a számát, a többiektől azt kérik, hogy önként lépjenek vissza ettől a megtiszteltetéstől. Úgy vélik, hogy a tudományos cikkek szerzőinek egy része "rejtélyes közreműködő", már ami a cikk létrejöttében játszott szerepüket illeti. Magát a cikket 35 szerző jegyzi, akik közül – a cikk 1. táblázata szerint – 31 "rejtélyes közreműködő". Őket a tényleges szerzők kérték fel társszerzőnek, és ők ezt elfogadták, a szerzők szerint többen úgy, hogy nem is voltak vele egészen tisztában, mit is vállalnak. Az egyik "rejtélyes közreműködő" visszalépett, amiért a szerzők külön lábjegyzetben mondtak neki köszönetet.

És ami nem vicc: a cikkel kapcsolatos keresgélés során bukkantam rá arra, hogy az egyik "rejtélyes közreműködő" az egyetemi honlapján (a pontos forrást ez alkalommal tapintatból nem adom meg) ezzel a cikkel publikációs listájának teljes értékű elemeként büszkélkedik.

### **3.2. Publikációk, publikációtípusok**

Sok évtizeddel ezelőtt a kutatóintézmények által beküldött publikációjegyzékek alapján készítettük első országos szintű tudományometriai felméréseink egyikét. Megdöbbentő volt az a heterogenitás, ami a beküldött listákból kitetszett. Volt intézmény, ahonnan lényegében csak a nemzetközi folyóiratokban megjelent cikkek adatait kaptuk meg, másokból viszont a munkatársak közreműködésével készült magazincikkek, rádióinterjúk, vagy éppen sehol sem publikált beszédek is a listára kerültek.

A helyzet az eltelt évtizedek alatt aligha változott drámaian, ahogyan azt az MTMT (Magyar Tudományos Művek Tára) munkatársai tanúsíthatják.

Hol vannak a tudományos publikálás határai? Mi „számít” tudományos publikációnak, és mi nem?

Ez az a kérdés, ahol először hivatkoznék a tudományometriai elemzések szerintem legfontosabb vezérelvére:

#### **VEZÉRELV**

**Az eszközök, módszerek alkalmassága mindig az elérendő céloktól függ.**

Valójában ez az elv a tudományometrián kívül is – mondhatjuk, bárhol – megállja a helyét, és ezt az elvet a tudományometriában és azon kívül is túlságosan gyakran figyelmen kívül hagyják. A svájci bicska *nem* olyan eszköz, amelyik minden feladatra alkalmas, hanem egy olyan *eszköztár*, amelyik nagyon sokféle feladathoz tartalmaz alkalmas eszközt. Hogy éppen melyik eszközt, az ollót vagy a fogpiszkálót használjuk, azt az adott feladat, az *elérendő cél* határozza meg.



Visszatérve az eredeti kérdésre: a tudományos publikációknak véleményem szerint nincs és nem is lehet egyetemesen használható meghatározása. Ezt a meghatározást mindig annak a feladatnak a függvényében kell megalkotnunk, amire készülünk. Természetesen felhasználhatunk mások által, esetleg más célokra felállított kritériumokat, de mindig alaposan meg kell gondolnunk, hogy azok az adott céljainknak megfelelnek-e. A kritériumok között lehetnek tartalmiak (pl. kísérleti eredményeken vagy elméleti levezetések alapján), hitelességet illetők (pl. két vagy több pozitív szakmai bírálat, jól dokumentált szakirodalom), formaiak (pl. jól meghatározott kérdésfeltevés, reprodukálható módszertan, egyértelmű következtetések) és sok minden más. Az MTA Könyvtárában működő MTMT – helyesen – azt az elvet követi, hogy minden tételt nyilvántart, amelyet a szerzője arra alkalmasnak tart, és olyan címkékké látja el őket, amelyeknek alapján a felhasználó eldöntheti, hogy egy adott szempontból relevánsnak tartja-e őket vagy sem.

#### KITÉRŐ

Annak eldöntése, hogy egy publikáció „tudományosnak” számít-e vagy sem, számomra ahhoz hasonlít, mint egy tevékenységformáról eldönteni, hogy „sport”-e vagy sem. Évtizedekig vitatéma volt, hogy sport-e a sakk, mostanában pedig az, hogy ha a sakk igen, akkor a póker miért nem. A szerencse szerepe nem feltétlenül kritérium, hiszen az számos sportban is lényeges tényező. Nem is olyan régen egészen abszurdnak hangzott, hogy a BMX freestyle vagy a break-tánc hivatalos sport legyen, ma pedig már az olimpiai sportágak között találjuk őket.

A legfontosabb kérdés újra meg újra az, hogy milyen célt akarunk elérni az ide vagy oda való besorolással. Az „elvi megfontolások” többnyire csak ürügyül szolgálnak. A valóságos kérdés az, hogy milyen fórumoktól, intézményektől várhatnak elismerést az egyes tevékenységek művelői – akár anyagi, akár bármilyen más formában. Ezt illetően pedig kizárólag az elismerést nyújtó intézmény jogosult dönteni. Alfred Nobel úgy döntött, hogy az örökségéből létrehozott díjjal matematikai kutatásokat nem jutalmaznak. Ezt semmiféle elvi megfontolás nem írhatja felül. Természetesen Nobel sohasem állította azt, hogy a matematikai kutatások nem tartoznak a tudomány körébe. Ez egyáltalán nem elvi, hanem egy nagyon is gyakorlati döntés volt.

Egy kiváló műfordító a közelmúltban azon kesergett, hogy mégoly kiváló műfordításait az MTMT nem sorolja a tudományos publikációk közé. Nyilván ezt azért sérelmezi, mert ő tudományos elismerésre vágyik. A döntés nyilván nem az MTMT és nem is a közvélemény, hanem az illető tudományos elismerést odaítélő grémium feladata. Ha ők az MTMT-beli besorolást tekintik irányadónak, az nem éppen szerencsés. Egyébként nem hiszem, hogy Albert Einstein valaha is irodalmi Nobel-díjra vagy olimpiai aranyéremre vágyott volna (vagy legalábbis nem a fizikában elért eredményeiért).

### ***3.3. Hivatkozások, idézetek***

A tudományos publikációk nélkülözhetetlen tartozéka az irodalomjegyzék, illetve a bennük szereplő hivatkozások. Formailag ezek lehetnek szövegközi említések, lapalji lábjegyzetek vagy a publikáció végén összegyűjtött lista tételei, a funkciójuk ugyanaz: megadni azokat a publikált előzményeket, amelyen az adott publikáció gondolatai, módszerei, eredményei alapulnak. A hivatkozások kiválasztása legalább olyan súlyos etikai problémákat vet fel, mint a szerzőké. Ideális esetben azok és csakis azok a tételek kerülhetnének a

hivatkozások közé, amelyeknek a publikáció létrehozásában tényleges és lényeges szerepük volt. A valóságban azonban szinte mindig tetten érhető valamiféle részrehajlás: a szerző személyének, intézményének felülreprezentálása, a konkurensok mellőzése, vélt vagy valódi tekintélyeknek a szerző elképzelései mögé állítása.

Az egyes hivatkozások tényleges súlyának megbecslése még a szerzői szerepek megítélésénél is nehezebb. Érdemi megítélést csak a hivatkozó és az összes hivatkozott cikk alapos összevetésével lehet elképzelni; ezt több tucat hivatkozás esetében a legelszántabb bírálótól, szerkesztőtől vagy elemzőtől sem lehet elvárni. Egyre szaporodnak azok a mesterséges intelligencián alapuló eljárások, amelyek a hivatkozások relevanciáját a hivatkozó cikkben elfoglalt helyzetük vagy a szöveggörnyezetük alapján próbálják megbecsülni. Érdekes próbálkozások ezek, amelyek gyakorlati értéke még meglehetősen kétséges.

A hivatkozások szerepét gyökeresen megváltoztatta az idézetelemzés megjelenése. Bár Eugene Garfield a Science Citation Index létrehozásával eredetileg a tudományos gondolatok visszakövethetőségét próbálta intézményesíteni (a korábban említett Shepard's Citations alapvető jogi forrásanyag mintájára), adatbázisa hamarosan a kutatásértékelés megkerülhetetlen, ezerszer áldott és átkozott eszköze lett. Minél többször hivatkoznak egy tudományos munkára, annál értékesebb – ez a kézenfekvőnek látszó, de valójában rendkívül ingatag állítás lett a kiindulópontja az idézetelemzés viharának, amely máig sem csitult el.

## KITÉRŐ

Az idézési indexek első példái a rabbinikus irodalomban lelhetők fel. Ezekben a XII. századtól kezdve összeállított gyűjteményekben a Biblia egyes paragrafusaira a magyarázó szövegekben előforduló hivatkozásokat gyűjtötték össze. Az angolszász precedensjog (esetjog) nélkülözhetlenné tette a jogi esetek indexelését, ami Angliában a XVIII. században kezdődött. Az Egyesült Államokban a XIX. század második felében jelentek meg ezek az indexek. Elsőként Kalifornia, majd New York állam eseteiről készültek ilyen gyűjtemények.

Frank Shepard (1848–1902) egy chicagói jogi kiadó ügynökeként látott meg üzleti lehetőséget az indexkötetek rendszeres kiadásában. 1873-ban megalapított cége ténylegesen 1875-ben kezdte meg tevékenységét Illinois állam eseteinek feldolgozásával. A Shepard's Citations hamarosan az az Egyesült Államok legjelentősebb jogi információforrásainak egyike lett. Mára az index online változatának, a Shepard's Citation Service-nek a használata úgyszólván kizárólagos, a nyomtatott köteteket csak hagyományból adják ki. A szolgáltatást jelenleg a LexisNexis, a RELX (korábban Reed-Elsevier) kiadói komplexum tagja nyújtja.

A hivatkozások értékmérő szerepének megjelenése természetesen erősen befolyásolta a szerzők hivatkozási stratégiáit, felerősítve a már korábban is meglévő részrehajlás tendenciáit, és újabb, etikailag egyre súlyosabban kifogásolható módszerekkel bővítve azt (kölcsonös hivatkozási klikkek, kikényszerített hivatkozások stb.).

Ezen a ponton térnék ki egy szóhasználati kérdésre. Ahogyan az angol nyelvű szakirodalom különbséget tesz a „reference” és a „citation” szavak között, a magyar szaknyelvben is célszerű megkülönböztetni a hivatkozás és az idézet

fogalmát. Röviden: hivatkozás az, amit adunk, idézet az, amit kapunk. A publikációk végén a hivatkozásokat találjuk, a publikációra kapott idézeteket más publikációk hivatkozásjegyzékeiből gyűjtjük össze. Természetesen minden egyes hivatkozás egyszersmind idézet is (a hivatkozott cikké), és megfordítva. A különbség csakis a nézőpontban van. Hasonlíthatjuk ezt a kiadás és a bevétel fogalmaihoz: minden egyes kiadás valaki más számára bevétel, és a kiadások és a bevételek összesített értéke minden pillanatban megegyezik.

A javasolt szóhasználat egyértelmű és követhető, nehéz olyan példát találni, ahol érdemes lenne eltérni tőle. A magyar nyelvben külön színezi a kérdést az, hogy az „idézni” szó a szó szerinti idézést (angolul „quotation”) is jelenti; a pontos szóhasználat ezt is egyértelműsítheti. „A cikkben idézték a definíciómra”: ez betű szerinti idézetet sejtet, „hivatkoztak a definíciómra”: feltüntették a forrást az irodalomjegyzékben. A bírósági *idézéseket*, és a szellem*idézést* most nem hoznám szóba.

### **3.4. A dokumentumok használatának mérése**

A nyilvános könyvtárakban rendszeres gyakorlat a helyben olvasási és kölcsönzési kérések statisztikai feldolgozása. Természetesen egy könyvtári tétel kikérése vagy akár kikölcsönzése nem igazán ad érdemi információt a tényleges használatáról: az olvasással töltött időről vagy a felhasznált információ mennyiségéről. Mindazonáltal az ilyen adatok hasznosak lehetnek a tárolási logisztika kialakításában, vagy a beszerzések megtervezésében.

A tudományos publikációk könyvtári használatára vonatkozó adatok még bizonytalanabbak. Folyóiratok esetében általában azt sem lehet tudni, hogy melyik cikk kedvéért kér ki az olvasó egy adott kötetet. A régi időkben a különlenyomatkéresek vagy a fénymásolatok kérésének száma adhatott némi

tájékoztatást egy-egy cikk iránti érdeklődésről. Az internet elterjedésével a kiadók a weboldalukról tudják regisztrálni az elérések, illetve a letöltések számát. Ezeknek a számoknak a relevanciája meglehetősen kérdéses (nem kis mértékben a cikkek elérésének és letöltésének számos alternatív lehetősége miatt), de jobb híján ezek az adatok szolgálnak a nagy bibliográfiai adatbázisokban (Web of Science, Scopus) a felhasználási (usage) mutatószámok alapjául.

Könyvek (akár tudományos, akár nem) esetében az eladási statisztikák (legismertebb az Amazon Best Sellers rangsora) jelezhetik a mű iránti érdeklődést.

### ***3.5. A közösségi média szerepe a tudományos kommunikációban***

A közösségi média platformjai az utóbbi évtizedek kommunikációs világát gyökeresen alakították át. Bár a tudományos kommunikáció részaránya ezeken a platformokon elenyésző, és a tudományos kommunikáción belül sem vált túlnyomóvá a közösségi médiában való jelenlét, ez a részarány évről évre érzékelhetően növekszik. Egyre nagyobb a külső és a belső készítés, hogy a kutatók eredményeiket a hagyományos publikációs csatornák mellett (rosszabb esetben helyettük) a közösségi média (főként a Facebook, a Twitter, a LinkedIn és különböző blogszolgáltatók) platformjain is nyilvánosságra hozzák. Ezek mellett megjelentek a tudományos világ saját közösségi médiái, amilyen a ResearchGate vagy az Academia.edu.

Bár több, mint kérdéses, hogy az ezeken a platformokon való jelenlét mértéke mennyire korrelál a kutatások tudományos értékével, a kutatásokban, illetve az eredmények kommunikációjában érdekelt, részben vagy egészben üzleti alapokon működő intézmények (egyetemek, kiadók, finanszírozó ügynökségek) gyakorlatilag megkövetelik, és folyamatosan monitorozzák a kutatási eredmények minél szélesebb körű közzétételét.

Hamar létrejöttek azok a szolgáltatók, amelyek ezt a monitorozást folyamatosan végzik, és eredményeiket a megrendelők rendelkezésére bocsátják.

A már korábban említett Altmetric.com 19 különféle közösségi megjelenési forma alapján képez egy kombinált mutatószámot. Az egyes megjelenési formák különböző súlyfaktoralal számítanak bele az összértékbe. A „legértékesebbek”: a hírekben való megjelenés (8 pont), a blogposzt (5), a szakpolitikai dokumentumok (3), a szabadalmak (3) és a Wikipédia (3). A Twitter 1, a LinkedIn 0.5, a Facebook és a YouTube 0.25 ponttal járul hozzá az összértékhez.

A PlumX mérőszámának pontos algoritmususa nem nyilvános. A Scopus előfizetői számára elérhető ez az érték, amelyet az egyes publikációk az idézetek (pl. szakirodalmi, szabadalmi), a használat (pl. kattintások, letöltések), a megjelölések (pl. kedvencek, könyvjelzők), az említések (pl. blogok, kommentek) és a közösségi médiában (pl. Facebook, Twitter) való jelenlét alapján kapnak.

A ResearchGate a regisztrált felhasználóira kiszámol egy RG-mutatót (RG score). Ennek meghatározása még a PlumX értékszámánál is átláthatatlanabb, és ezért heves kritikák forrása. Valójában az RG-mutatóhoz még az sem szükséges, hogy a regisztrált személynek akár egyetlen publikációja is legyen; a kommentelés, a vitákban való részvétel, a nagy számú követő összegyűjtése elegendő a kiemelkedő RG-mutató eléréséhez. Sokan mindennek ellenére az RG-mutatót a kutató kommunikációs képességét és ezáltal a közösségre gyakorolt hatását jól jellemző indikátornak tartják.

### **3.6. Statisztikai alapok**

A statisztika két alapvető művelete az osztályozás (csoportosítás) és a számlálás. Mondhatjuk, hogy ez a statisztikai vizsgálatok két kreatív eleme, minden más csak számszaki műveletek sora, többé-kevésbé egyszerű matematikai manipulációk alkalmazása. Az osztályozás és a számlálás azonban millió buktatót rejt, és ahogy korábban szó volt róla, nincsenek „jó” és „rossz” módszerek, a választás mindig az elérendő céltól függ.

Tekintsük példának azt, hogy egy könyvtár könyveiről kell statisztikát készíteni. A könyveket nyilván temérdekféleképpen lehet csoportosítani. Lehet szín szerint, nagyság szerint, szerző szerint, ISBN szám szerint. A nyilvános könyvtárak könyvei esetében könnyen osztályozhatjuk a könyveket használati (olvasási, kölcsönzési) gyakoriság alapján. Az osztályozás szempontja (ismérv) lehet mennyiségi (pl. méret) vagy minőségi (pl. szín). Egyes ismérvek szerint (pl. nagyság, szerzői abc-rend) szerint sorba rendezni is lehet, mások (pl. szín) szerint csak osztályokba sorolni. Számtalan probléma akadhat már az osztályozás során is. A könyvnek melyik méretét vegyük figyelembe? Milyen színek kategóriákat használjunk? Hová soroljuk a többszínű könyveket? Hasonlóképpen csapdákkal teli a számlálás művelete is. Hogyan számláljuk a többkötetes könyveket? A közös kötetben megjelent (vagy utólag egybekötött) könyveket? A több példányban, több kiadásban előforduló könyveket?

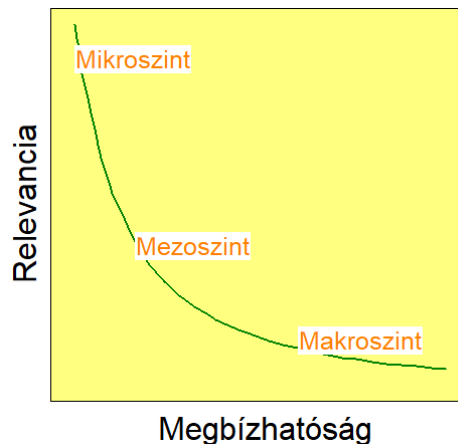
A tudománymetriai objektumok osztályozását végezhetjük más tudománymetriai ismérvek vagy azon kívüli kritériumok szerint. A szerzőket csoportosíthatjuk publikációik száma, a publikációkat társszerzőik száma szerint; a publikációkat osztályozhatjuk akár hivatkozásaik száma (az irodalomjegyzékük hosszúsága), akár a kapott idézeteik száma szerint (itt láthatjuk például, hogy milyen fontos különbséget tenni a két fogalom – hivatkozás és idézet – között).



A szerzők, publikációk, hivatkozások, idézetek csoportosítását gyakran olyan szempontok szerint végezzük, amelyek besorolhatók a „nagyságrendi szintek” valamelyikébe:

- a *makroszint* a nagyobb földrajzi egységeket (pl. országok, régiók) vagy a tudomány nagyobb területeit (pl. fizika, kémia, társadalomtudomány) jelenti;
- a *mezoszint* az intézmények, szűkebb szakterületek (pl. funkcionálmélet, génebézészet), alkalmasint a folyóiratok szintje;
- a *mikroszint* az egyes kutatók, esetleg kutatócsoportok (műhelyek) vizsgálatával foglalkozik.

Érdemes szem előtt tartani, hogy a három szinten végzett vizsgálatok két fontos sajátossága: a relevanciájuk és a megbízhatóságuk reciprok viszonyban állnak egymással (20. ábra).



20. ábra Relevancia és megbízhatóság

### **3.7. Statisztikai mutatószámok**

A statisztikai elemzés előkészítése azzal fejeződik be, hogy a vizsgált sokaság elemeit csoportokba rendeztük, és a csoportok elemeit megszámláltuk. Bízunk benne, hogy az osztályozást és a számlálást a vizsgálat céljának legmegfelelőbb módon végeztük el.

A következő lépés az, hogy ebből a számhalmazból egy vagy több olyan jellemző számértéket, *statisztikai mutatószámot* képezzünk, amelyek – újra és újra hangsúlyozzuk – *a vizsgálat céljainak megfelelő módon* jellemzik a vizsgált sokaságot. Semmiképpen nem beszélhetünk tehát jó vagy rossz mutatószámokról, viszont nagyon is beszélhetünk a mutatószámok helyes és helytelen használatáról.

A mennyiségi ismérvek statisztikai mutatószámainak alapvető típusai a helyparaméterek és a szóródási paraméterek.

A helyparaméterek a vizsgált mennyiségi jellemző jellemző „közepes” értékét képviselik. Ide tartoznak a különféle átlagértékek vagy a medián. Az átlagértékek közül a számtani, a mértani és a harmonikus átlag fogalma már (vagy még) a középiskolából ismerős lehet; a gyakorlatban még a különféle súlyozott átlagok kaphatnak fontos szerepet. (Például egy nemzetközi összehasonlításban a vizsgált országhoz földrajzilag közelebb fekvő országok értékeit nagyobb súllyal vehetjük figyelembe, mint a távolabbiakét.) A medián az az érték, amelyik a sorba rendezett sokaságot pontosan két egyenlő részre osztja; pont azonos számú nála kisebb és nagyobb elem van a sokaságban. A kvartilisek a mediánnal együtt a sokaságot négy egyenlő részre osztják: a felső kvartilis a medián fölötti elemek mediánja, az alsó kvartilis a medián alatti elemek mediánja. Több egyenlő számú részre való osztással különféle percentiliseket definiálhatunk.

A szóródási paraméterek legegyszerűbb példája a (minta)terjedelem: a legnagyobb és a legkisebb érték különbsége. A számtani átlagérték mellé szóródási paraméternek az átlagos négyzetes középeltérést (szórás) szokás megadni, míg a mediánhoz tartozó szóródási paraméter az interkvartilis tartomány (a felső és alsó kvartilis különbsége).

A statisztikai mutatószámok sajátos fajtái az extrémérték statisztikák, amelyek az előforduló legkisebb és legnagyobb értékeket jellemzik. Ilyeneknek nagy jelentőségük van pl. a sportstatisztikákban (rekordok) vagy a vízállásértékek elemzésében (árvizek). Érdekes módon ezzel a kategóriával rokon a tudományometriában az elmúlt évtizedekben bevezetett Hirsch-féle h-index.

Érdemes szem előtt tartani, hogy a matematikai statisztika több évszázados múltja során a statisztikai mutatószámok ezreit dolgozták ki és vizsgálták meg. Aligha képzelhető el olyan gyakorlati probléma, amelyhez legjobban illeszkedő mutatószámokat ne lehetne megtalálni a szakirodalomban. A helyes választáshoz azonban ennek az óriási irodalomnak az alapos ismerete szükséges, ezért – az ésszerű önművelés mellett – esetenként nem árt statisztikus szakértő segítségét is igénybe venni.

#### KITÉRŐ

A statisztikáról szóló legrövidebb összefoglalásból sem hiányozhat néhány azok közül a – többnyire gunyoros – elmésségek közül, amelyeket a témáról a közvélekedés jellemzőnek talál. A mondások szerzőinek személye általában bizonytalan, az *idézetek* nincsenek alátámasztva *hivatkozásokkal*.

A legtöbb ember úgy használja a statisztikákat, ahogyan a részeg ember a lámpaoszlopot: inkább támaszként, mint megvilágításra. (Mark Twain)

A statisztikák olyanok, mint a bikini. Amit elárulnak, az szuggesztív, de a lényegét elrejtik. (Aaron Levenstein)

Ha mérésed statisztikára szorul, akkor mérd jobban! (Ernest Rutherford)

A statisztika alapvető jellemzője a részletek körültekintő és rendszeres figyelmen kívül hagyása. (Erwin Schrödinger)

## 4. A tudománymetria használatának gyakorlati példái

A történeti és elvi alapok áttekintése után rátérhetünk a tudománymetria gyakorlati alkalmazásának kérdéseire. Egy gyakorlati útmutató voltaképpen egy szakácskönyv. Legfontosabb elemei: a nyersanyagok, az elkészítési (előkészítési, főzési, sütési) módszerek és a tálalás. Esetünkben a nyersanyagok a bibliográfiai adatforrások, illetve az azokból kinyerhető nyers adathalmazok. Az adatelőkészítés ugyanolyan fontos, mint a húsok megfelelő pácolása, vagy a tészta megkelesztése; a sütés-főzés pedig a mutatószámok meghatározására alkalmazott „receptek”, algoritmusok használata. Végül a tálalás a kész eredmények prezentációja.

Az, hogy egy szakácskönyvből melyik ételt választjuk elkészíteni, az természetesen függ a rendelkezésre álló nyersanyagoktól, de a nyersanyagok és a receptek kiválasztása is mindenekelőtt a sütés-főzés *céljától* függ: gyors és tápláló enivalót szeretnénk magunknak, látványos fogásokat a vendégeinknek, vagy éppen speciális diétás ételt egy lábadozónak.

Talán nem felesleges, ha ezen a helyen megismételjük a tudománymetriai elemzések legfőbb vezérelvét:

### VEZÉRELV

**Az eszközök, módszerek alkalmassága mindig az elérendő céloktól függ.**

A szakácskönyvek között vannak olyanok, amelyek, ha nem is teljességre, de arra törekszenek, hogy az étlap minden elemére (előételek, levesek, főételek, desszertek) bőséges választékot bocsássanak rendelkezésre. Mások csak egy-egy speciális témára (pl. főzelékek) fókuszálnak. Az itt következő áttekintés –

mondjuk úgy – egy kissé esetlegesen összeválogatott ízelítő a tudománymetria szakácskönyvének választékából.

#### ***4.1. A tudománymetria főbb adatforrásai***

A tudománymetria hőskorában az elemzések adatforrásai kézzel összeállított listák voltak. Lotka klasszikussá vált munkája egy fizikatörténeti könyv (nyilván kézzel összeállított) névmutatóját használta adatforrásul. Az MTA Könyvtárában annak idején a kutatóintézetek által beküldött, írógéppel írt publikációjegyzékekből készítettük első elemzéseinket. Ilyen ma már legfeljebb egészen kivételes esetekben fordulhat elő.

A tudománymetriai vizsgálatok döntő többsége a számítógépes (legtöbbször interneten elérhető) bibliográfiai és idézettségi adatbázisokat használja adatforrásul. Ezeknek az adatbázisoknak egy része előfizetéses alapon működik. Ezek legismertebb képviselői a Clarivate tulajdonában lévő Web of Science ([apps.webofknowledge.com](https://apps.webofknowledge.com)) és az Elsevier által működtetett Scopus ([www.scopus.com](http://www.scopus.com)). Mindkét adatbázis a természettudományok, társadalomtudományok és a humán tudományok nagy részét lefedi. Ezek az adatbázisok meglehetősen drágák, de a magyar kutatók számára az Elektronikus Információs szolgáltatás Nemzeti Program (EISz) keretében általában elérhetők. A drága előfizetés fejében az adatbázisok bizonyos mértékben ellenőrzött, megtisztított és további feldolgozásra alkalmas adatokat szolgáltatnak.

Az 1999-ben létrehozott Crossref ([www.crossref.org](http://www.crossref.org)) a tudományos publikációk metaadatainak egyik legteljesebb gyűjtőhelye. A Crossref legfőbb funkciója a publikációk egyedi azonosítóinak (Digital Object Identifier; DOI) kiosztása és nyilvántartása. Az adatbázis jelenleg (2021. december) több, mint 120 millió tétel adatait tartalmazza. A metaadatokat az adatszolgáltatók

(kiadók) jóváhagyásától függően tárolják. Miután 2020. decembere óta a legnagyobb tudományos kiadó, az Elsevier, is hozzájárult a cikkek hivatkozásainak metaadatként való nyilvántartásához, a jövőben a Crossref a világ legteljesebb publikációs és idézettségi adatforrásává válhat. Az adatbázis teljesen nyitott, teljes adatállománya (több terabájtnyi adat) szabadon hozzáférhető.

A többi ingyenes nagy bibliográfiai adatbázis mind valamelyik nagy informatikai cég tartozékaként működik. A Google Scholar (GS) ([scholar.google.com](http://scholar.google.com)) és a Microsoft Academic (MA) ([academic.microsoft.com](http://academic.microsoft.com)) tulajdonosát a nevük is elárulja. (Megjegyezzük, hogy az MA 2021 decemberétől megszűnt.) A Dimensions ([www.dimensions.ai](http://www.dimensions.ai)) a Digital Science, a Meta ([www.meta.org](http://www.meta.org)) 2017 óta a Chan Zuckerberg Initiative (CZI) tulajdonában van.

Ez a legutóbbi, a CZI 2021 októberi bejelentése szerint 2022 márciusától beszünteti működését. Érdekesség, hogy ezzel egyidőben jelentették be, hogy a „Zuckerberg-birodalom” korábban Facebook, Inc. néven ismert anyacége a továbbiakban Meta néven működik tovább.

Ezeknek az adatforrásoknak a legfőbb vonzereje az ingyenesség; ennek ára az, hogy adatok helyenként gondozatlanok és további feldolgozásra kevésbé alkalmasak.

Ezeken az általános gyűjtőkörű adatforrásokon kívül számos szakterületnek van saját bibliográfiai adatbázisa.

A szakterületi adatbázisok közül a legjelentősebb a Medline, az Egyesült Államok Orvosi Könyvtárának bibliográfiai adatbázisa ([www.nlm.nih.gov/medline](http://www.nlm.nih.gov/medline)). Az 1964-ben alapított adatbázis jelenleg kb. 30 millió élettudományi (főként orvosbiológiai) publikáció metaadatait tartalmazza.

A CiteSeerX a tudományos cikkek keresőmotorja és digitális könyvtára főként az információ- és számítástudomány területén. Ők szabadalmaztatták először a tudományos cikkek és idézetek internetről való összegyűjtésének és indexelésének a módszerét. Mintájára több más területen is próbálkoztak hasonló rendszerek létrehozásával (eBizSearch, ChemXSeer, ArchSeer), de kevesebb sikerrel.

A számítástudomány egyik legteljesebb bibliográfiája a DBLP adatbázisban található. A DLBP eredetileg a "DataBase systems and Logic Programming" rövidítése volt. Később a "Digital Bibliography & Library Project" nevet illesztették rá. Jelenlegi hivatalos neve: "The DBLP Computer Science Bibliography". A Trieri Egyetem indította el 1993-ban; jelenleg a Dagstuhl Kastélyban létrehozott Leibniz Informatikai Központ (Schloss Dagstuhl – Leibniz-Zentrum für Informatik) működteti. A folyóiratcikkekén kívül a konferenciakiadványok is képviselve vannak benne, amelyek a számítástudomány területén a szakirodalom jelentős részét alkotják. A nyílt hozzáférésű publikációk teljes szövege, a többinek csak metaadatai hozzáférhetők.

A szakterületeken kívül egyes országok és régiók is létrehoztak bibliográfiai adatbázisokat. Magyarországon ennek legfőbb képviselője az MTA által kezdeményezett, 2009-ben létrehozott Magyar Tudományos Művek Tára (MTMT; [www.mtmt.hu](http://www.mtmt.hu)). A SCIELO (Scientific Electronic Library Online; [scielo.org](http://scielo.org)) a fejlődő országok számára 1997-ben Brazíliában alapított tudományos bibliográfiai adatbázis. Létrehozásának célja a nagy nemzetközi adatbázisok angol nyelvű túlsúlyának kiegyensúlyozása volt a spanyol és portugál nyelvű szakirodalommal. Jelenleg 16 (főleg latin-amerikai) ország tartozik a hálózatához.

A nagy kommerciális adatbázisok (WoS, Scopus) beépített lehetőséget nyújtanak egyszerű statisztikák készítésére. A WoS adatbázis „Analyze”

opciója számos szempont szerint készít statisztikát a keresés eredményeként kapott publikációkról. A 21. ábrán bemutatott példán a publikáló országok szerinti statisztika látható, de megtalálhatjuk rajta a többi lehetséges szempontot is.

A PubMed kizárólag a publikálási év szerinti statisztikát készíti el a felhasználói felületén (22. ábra). Minden mást a keresési eredmények letöltése után, saját eszközökkel (ami egyszerűbb esetekben lehet akár Excel is) lehet elkészíteni.

Az ingyenes adatbázisok folyamatosan fejlesztik a felhasználók hozzáférési lehetőségeit a nyers adatokhoz, ill. egyszerű összesített statisztikákhoz (pl. toplisták), de a lehetőségek korlátozottak. A legnagyobb probléma az, hogy az adatok pontossága, megbízhatósága mérsékelt, ezért a belőlük készített elemzések is nagyon korlátozott érvényességűek.

A leggyakrabban használt adatbázisokból egységes tartalmú és formájú statisztikákat lehet készíteni a holland születésű, jelenleg Angliában élő Ann-Wil Harzing Publish or Perish ([harzing.com/resources/publish-or-perish](http://harzing.com/resources/publish-or-perish)) programjával. Az internetről különféle számítógépes platformokra letölthető szabad hozzáférésű szoftver egységes felhasználói felülettel enged hozzáférést az említett adatbázisokhoz. (Természetesen az előfizetéses adatbázisok esetében a felhasználónak rendelkeznie kell a hozzáférési jogosultsággal.) Az adatbázis kiválasztása után össze lehet gyűjteni a vizsgálandó publikációhalmazt, majd ezekre különféle tudományometriai mutatószámok szerinti rangsorokat és összesített statisztikákat lehet meghatározni (23-24. ábra).

Sohase felejtjük el, hogy a legkiválóbb szoftverrel előállított mutatószámok megbízhatósága sem lesz nagyobb az előállításukhoz felhasznált adatok megbízhatóságánál, valamint, hogy az ezekkel a szoftverekkel végzett



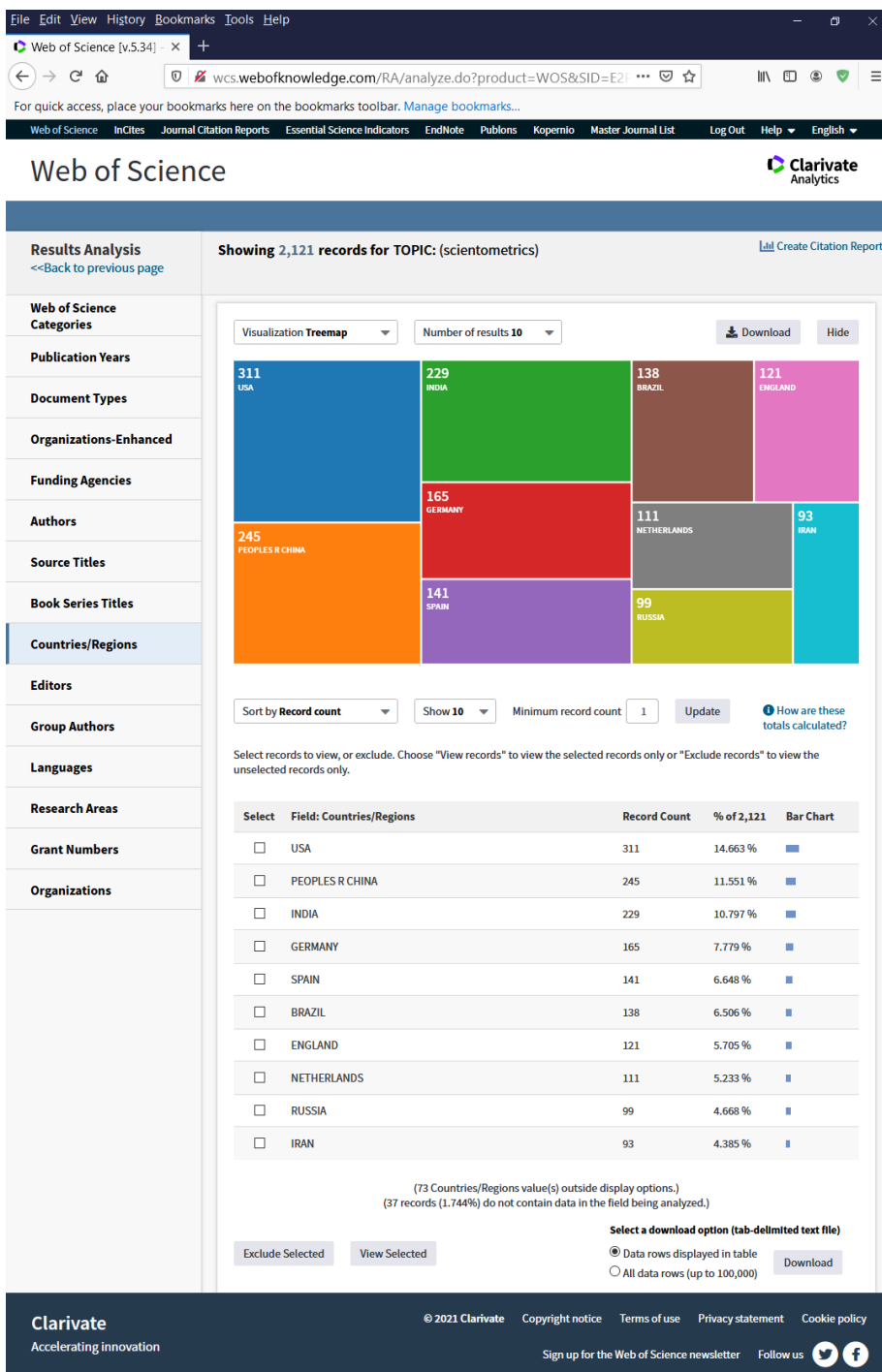
elemzések legfeljebb előzetes tájékozódásra alkalmasak és nem érdemi döntéseket befolyásoló iránymutatásra. Mondjuk úgy: félkész ételek.

### KITÉRŐ

A „modernista szakácsművészet pápája”, Nathan Myhrvold ezt írja: „Jó ételeket csak jó alapanyagokból lehet készíteni. Volt idő, amikor ezt az alapigazságot csak a legkiválóbb szakácsok méltányolták. Ma már, úgy látszik, az egész élelmiszervilág egyre nagyobb hangsúlyt helyez az alapanyagok minőségére.”

Nagyon megfontolandó igazság. A legleleményesebb algoritmusok sem tudják orvosolni a kiinduló adatok fogyatékoságait. Az ingyenes adatbázisok adataiból készített elemzések legfeljebb közelítő eredményeket szolgáltathatnak.

De érdemes John W. Tukey intését is szem előtt tartani: „Sokkal jobb egy közelítő válasz egy akár bizonytalanul megfogalmazott, de helyes kérdésre, mint egy egzakt válasz egy precízen megfogalmazott, de helytelen kérdésre.”  
(Az adatelemzés jövője, 1962.)



21. ábra A „scientometrics” téma irodalmának országok szerinti megoszlása a WoS alapján

Web of Science [v.5.35] - scientometrics - Search

https://pubmed.ncbi.nlm.nih.gov/?term=scientometrics&timelin...

National Library of Medicine  
National Center for Biotechnology Information

PubMed.gov

scientometrics

Advanced Create alert Create RSS

Sorted by: Publication date

RESULTS BY YEAR 986 results

MY NCBI FILTERS

TEXT AVAILABILITY

- Abstract
- Free full text
- Full text

ARTICLE ATTRIBUTE

- Associated data

ARTICLE TYPE

- Books and Documents
- Clinical Trial
- Meta-Analysis
- Randomized Controlled Trial
- Review
- Systematic Review

PUBLICATION DATE

- 1 year
- 5 years
- 10 years
- Custom Range

Additional filters

Reset all filters

1  **Ulcerative Colitis: A Critical Approach to the Global Research Output Employing Density-equalizing Mapping and **Scientometric** Methods.**

Cite Schöffel N, Brüggmann D, Klingelhöfer D, Bendels MHK, Groneberg DA.  
J Clin Gastroenterol. 2021 Mar 1;55(3):e19-e26. doi: 10.1097/MCG.0000000000001351.  
Share PMID: 32324679

Hence, no current, detailed and comprehensive knowledge regarding the worldwide research architecture of UC has been established until now. METHODS: We conducted a **scientometric** study employing the previously validated NewQIS (New Quality and Quantity Indices in Science) p...

2  **Data integration enables global biodiversity synthesis.**

Cite Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D.  
Proc Natl Acad Sci U S A. 2021 Feb 9;118(6):e2018093118. doi: 10.1073/pnas.2018093118.  
Share PMID: 33526679

3  **Characterizing Highly Cited Papers in Mass Cytometry through H-Classics.**

Cite Di Zeo-Sánchez DE, Sánchez-Núñez P, Stephens C, Lucena MI.  
Biology (Basel). 2021 Feb 2;10(2):104. doi: 10.3390/biology10020104.  
Share PMID: 33540586

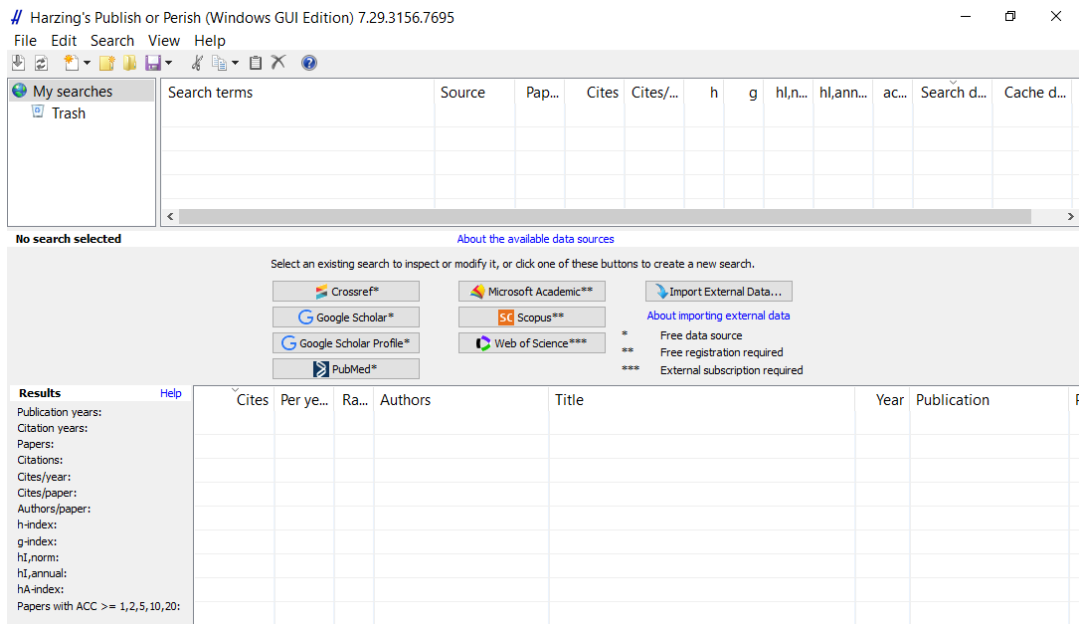
4  **Scientometric analyzing the output of researchers and organizations on COVID-19 for better conducting the scientific efforts: with a glance to endocrinology.**

Cite Atlasi R, Noroozi Chakoli A, Ramezani A, Tabatabaei-Malazy O, Larjani B.  
J Diabetes Metab Disord. 2021 Jan 29;1-12. doi: 10.1007/s40200-020-00718-7. Online ahead of print.  
Share PMID: 33532371 **Free PMC article.**

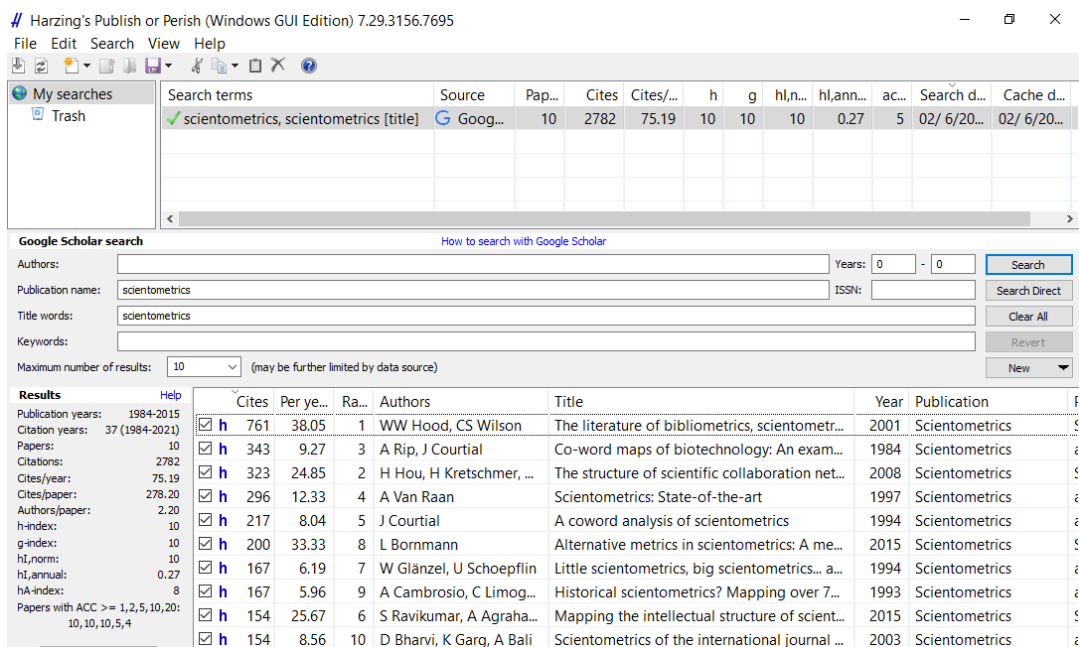
5  **Rethinking SME default prediction: a systematic literature review and future perspectives.**

Cite Ciampi F, Giannozzi A, Marzi G, Altman EI.  
Scientometrics. 2021 Jan 29;1-48. doi: 10.1007/s11192-020-03856-0. Online ahead of print.  
Share PMID: 33531720 **Free PMC article.**

22. ábra A „scientometrics” téma irodalmának publikálási évek szerinti megoszlása a PubMed alapján



23. ábra A Publish or Perish lekérdezési felülete



24. ábra Publish or Perish idézettségi lista a Google Scholar adatai alapján

## 4.2. Mutatószámok és használatuk

A bibliometriai mutatószámok félezer oldalas kézikönyvében [Todeschini & Baccini, 2016] körülbelül ugyanannyi mutatószám leírása szerepel. A meghatározások között vannak egyszerűek:

$$CL = Y_c - Y_P ,$$

és vannak bonyultabbak:

$$s^2 = \frac{1}{[A \cdot (A - 1)]^2 \cdot P \cdot k^2} \cdot \left[ \sum_{i=1}^A \sum_{j=1}^A \sqrt{\sum_{a=1}^q f_{ij}^{(a)}} \right]^2$$

Ha hozzávesszük, hogy 2016 óta a tudományometriai mutatószámokkal foglalkozó irodalom közel megkétszereződött, beláthatjuk, hogy ennek a témának a felületes áttekintése is messze meghaladja egy rövid előadás kereteit.

Néhány példát tekintünk meg tehát csupán, és ezek kapcsán teszünk néhány általánosabb észrevételt.

### ➤ Impakt faktor

Ahol a tudományetriáról egyáltalán szó esik, aligha kerülhető el az impakt faktor emlegetése. Mondanak róla forrót és hideget, indulatokat kavarr, és – valljuk be – sok-sok csacszkodás és értetlenség övezi [Schubert, 2015a].

Az impakt faktor definíciója (látszólag) egyszerű: egy folyóirat tavalyi és tavalyelőtti cikkeire idén kapott idézetek átlagos száma. A „hatástényező” (ahogy sokszor magyarul nevezik) tehát egy folyóirat kétéves időszakának az idézettségi hatását jellemzi. Ennyit: nem többet, és nem kevesebbet. Így definiálta Eugene Garfield 1972-ben [Garfield, 1972], és így található meg 1976 óta a Journal Citation Reports (JCR) kötetében, illetve adatbázisában.

Érdemes hangsúlyozni, hogy az impakt faktor mindenekelőtt egy üzleti (mellék)termék, ami a fő termék, a Science Citation Index használatának sokoldalúságát hivatott bemutatni (a JCR használata maga is külön előfizetéshez kötött), minden egyéb használata járulékos haszonnak – vagy helytelen használat esetén járulékos kárnak – tekinthető.

A gyakorlatban több, az eredetitől eltérő értelmezésben is használják az „impakt faktor” elnevezést. Maga a JCR is tartalmaz egy „5-éves impakt faktor” mutatószámot, ami az előző öt évben megjelent cikkek tárgyevi idézettségének az átlaga. Az irodalomban pedig se szeri se száma a különféle időtartamokra és publikációhalmazokra vonatkozó „impakt faktoroknak”. Ami mindegyikben közösnek tekinthető, az az, hogy a mutatószám egy bizonyos időtartamban (publikációs ablak) megjelent cikkeknek egy másik időtartamban (idézettségi ablak) kapott idézettségi átlaga.

Az impakt faktorról szembeni kritikák áttanulmányozásával összeállíthatjuk az ellenvetések egy általános tipológiáját.

- Az idézettségen vagy még általánosabban a tudományometriai módszereken alapuló értékelési eljárások elutasítása;
- Az átlag, mint egy sokaságot jellemző mutatószám alkalmasságának kétségbevonása;
- A publikációs és idézettségi időablakok bírálata;
- A szakterületi különbségek figyelmen kívül hagyása;
- A folyóirattípusok közötti különbségek figyelmen kívül hagyása (pl. review folyóiratok);
- A dokumentumtípusok közötti különbségek figyelmen kívül hagyása;
- Az önidézetek szerepének figyelmen kívül hagyása;
- Az impakt faktornak nem folyóiratok értékelésére való felhasználása (egyének, intézmények, stb.), ill. a felhasználás módjának megkérdőjelezése.

Azt mondhatjuk, hogy az ellenvetések nagy része nem teljesen megalapozatlan, de érvényességüket, relevanciájukat csak egyes esetekre, és mindenképpen a konkrét alkalmazási cél figyelembe vételével lehet megvizsgálni.

Az impakt faktor kutatásértékelési szerepével kapcsolatos kétségek három szinten fogalmazhatók meg.

- 1. Alkalmasak-e a tudománymetria módszerei (publikációk, idézetek statisztikai elemzése) arra, hogy a tudományos kutatás minőségét, színvonalát illető érdemi információkkal járuljon hozzá a kutatási teljesítmény értékeléséhez?
- 2. Egy folyóiratok jellemzésére kidolgozott mutatószámoknak van-e keresnivalója olyan esetekben, amikor az értékelés tárgya nem folyóirat, hanem pl. intézmény, kutatócsoport vagy esetleg egyén?
- 3. Az impakt faktor (Garfield-faktor) vagy valamelyik másik folyóirat-mutatószám alkalmasabb-e az értékelési feladatban való felhasználásra?

A három kérdés nyilvánvalóan hierarchikusan követi egymást. Ha valamelyik kérdésre a válasz „nem”, akkor a következő kérdés(ek) irrelevánsakká válnak.

Ad 1. Az első kérdés könyvtárnyi irodalmának taglalásába most nem bocsátkoznék. Érdekes megfigyelni, hogy a legkidolgozottabb, rendszeresen használt értékelési rendszerek (pl. az Egyesült Királyság vagy Ausztrália egyetemi értékelési rendszerei) az elmúlt évtizedekben bizonyos periodicitást mutattak a tudománymetriai módszerek elfogadását (sőt, néha túlértékelését), ill. elutasítását illetően. Az Egyesült Királyság ilyen rendszere, a Research Evaluation Framework 2014-es álláspontja szerint „az idézettségi információk nem kellőképpen robusztusak ahhoz, hogy képletszerűen vagy elsődleges mutatószámokként használjuk őket, de ezeknek az adatoknak

tájékoztató és segítő szerepük lehet a szakértői vélemények kialakításában”. Kellően semmitmondó, de nem egészen haszontalan álláspont.

Ad 2. Bár igaz az, hogy egy folyóiratban megjelent cikkek idézettsége határozza meg a folyóirat impakt faktorát, és nem fordítva, de vitathatatlan, hogy a kiemelkedő impakt faktorú folyóiratok a cikkek kiválasztási mechanizmusával, a bírálati rendszer szigorúságával igyekeznek fenntartani a megjelent cikkek iránti érdeklődést, ezen keresztül az idézettséget. Ezért aztán az ilyen folyóiratokban való megjelenés már eleve egy bizonyos elismerést jelent.

Természetesen egy bizonyos idő eltelte után egy cikknek már magának kell bizonyítani érdemeit. A tudománymetria keretei között ezt a kapott idézetekkel érheti el. Ennek megfelelően a folyóirat-mutatószámok (így az impakt faktor) a kutatásértékelésben kettős szerepet játszanak.

„Impakt faktor a számlálóban”: Ez a viszonylag friss (kb. az elmúlt 2–3 évben megjelent) cikkekre vonatkozik: az impakt faktor ilyenkor a publikációs stratégia ambíciózusságának, ill. a friss publikációk jövőben várható idézettségének megítélését szolgálja.

„Impakt faktor a nevezőben”: Ez a régebbi cikkekre vonatkozik: ilyenkor az impakt faktor az elvárt idézettség mércéjéül szolgál, amihez a ténylegesen kapott idézettséget hozzámérhetjük.

Egy nagy impakt faktorú folyóiratban megjelent cikk szerzője tehát a megjelenést követően egy ideig joggal élvezheti a megelőlegezett bizalmat, de azután az előleget vissza kell fizetni, a nagyobb impakt faktornak megfelelő nagyobb idézettséget kell elérni ahhoz, hogy teljesítse az elvárásokat.

Ad 3. Azt, hogy valamelyik mutatószám alkalmasabb-e egy másiknál, mindig csak egy konkrét értékelési feladat kapcsán lehet eldönteni (akkor sem mindig könnyű). Semmi ok nincs azt feltételezni, hogy ne legyen olyan feladat, amikor



az impakt faktor lenne a legalkalmasabb mutatószám, de bizonyára sok esetben nem az. Ezért kell egy minél szélesebb eszköztár, és nagy elemzői tapasztalat az értékelési feladatok eredményes elvégzéséhez.

A DORA néven elhíresült dokumentumot (*San Francisco Declaration on Research Assessment*; [sfdora.org](http://sfdora.org)) 155 egyéni és 82 intézményi aláíró hitelesítette az Amerikai Sejtbiológiai Társaság (ASCB) éves konferenciáján San Francisco-ban 2012 december 16-án. A későbbiek során a deklarációhoz több ezer egyéni és intézményi aláíró is csatlakozott. A kétségkívül alapvetően jó szándékú kezdeményezés hangsúlyosan fellép az impakt faktornak a kutatás értékelésében való helytelen használata ellen. A deklaráció szövegében szereplő használati mód – vagyis az, hogy az impakt faktort az egyes cikkek, ill. szerzőik közvetlen értékmérőjéül használjuk – ugyanis egyértelműen helytelen. Sajnálatosan – de nem egyedülálló módon – a jó szándékú intés a későbbiekben dogmává merevülve az impakt faktor (és részben a vele rokon mutatószámok) teljes diszkreditálásának eszközévé vált. Ez majdnem olyan káros, mint a kritikátlan elfogadás.

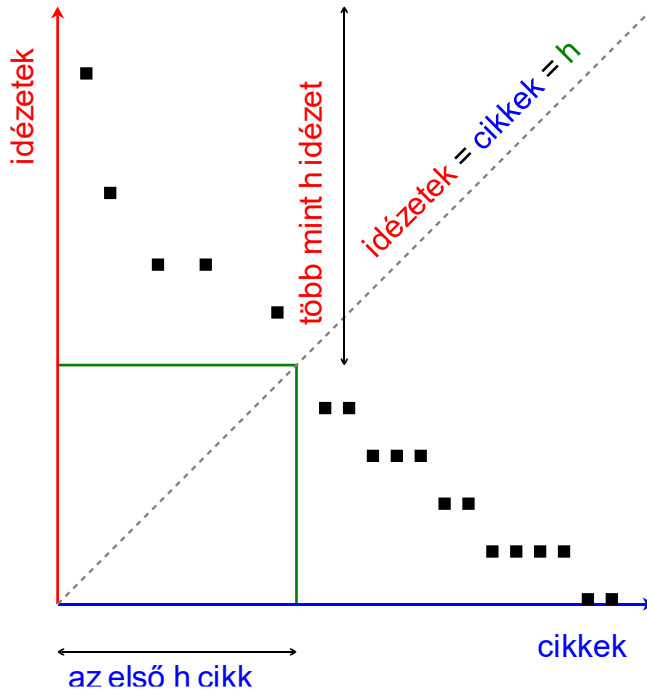
Láthatóan tisztában van ezzel a Németországi Tudományos Orvosi Társaságok Egyesülete (AWMF) is. Az orvosi kutatás teljesítményének értékeléséről kiadott brosúrájukban a DORA aláíróiként természetesen teljes egyetértésüket fejezik ki az abban foglaltakkal; hosszabb paragrafusokat szó szerint is átvesznek belőle. A tudományometriai mutatószámokat illetően azonban hozzátesszik: „Ezeknek a mutatószámoknak bizonyítékokon alapulónak, átláthatóknak és kivitelezhetőeknek kell lenni. A mutatószámok bizonyított tapasztalati háttére meg kell, hogy feleljen az értékelés céljának. Az egyszerűség nem lehet a fő kritérium, ez nem vezethet alkalmatlan eszközök használatához. [...] Módszertanilag adekvát értékelések készítéséhez szükség esetén tudományometriai szakértőket kell 'megvásárolni'. ”

Ennek ellenére számos kutatásértékeléssel foglalkozó intézmény (pl. az MTMT) radikálisan száműzte eszköztárából az impakt faktort, és más, nem kevésbé kétséges érvényességű mutatószámokat épített be rendszerébe.

➤ A h-index

Az impakt faktor, mint vitatéma méltó vetélytársra talált 2005-ben, amikor Jorge E. Hirsch, argentin születésű amerikai fizikaprofesszor közzétette javaslatát az egyéni kutatási termelékenység egy kvantitatív mutatószámára [Hirsch, 2005]. A h-index néven ismertté vált (sokszor Hirsch-indexnek is nevezett) mutatószám az impakt faktorhoz hasonlóan máig sem csituló viharokat kavart a tudománymetria világában és azon kívül is.

Míg azonban az impakt faktor matematikai-statisztikai értelmezése egyértelmű: egy idézeteloszlás átlagértéke, a h-index egy kevés előzménnyel rendelkező statisztikai mutatószám. A h-index meghatározásához a publikációkat a kapott idézeteik száma szerint sorba kell rendeznünk, és meg kell keresnünk azt a publikációt, amelynek a rangszáma (hányadik a sorban) megegyezik az értékével (hány idézetet kapott). Ez a szám lesz az adott idézeteloszlás (vagyis a publikációhalmaz) h-indexe (25. ábra).



25. ábra A Hirsch-féle h-index grafikus értelmezése

Látszólagos egyedisége ellenére a h-index elhelyezhető a hagyományos matematikai-statisztika rendszerében, sőt, egyszerű közelítő matematikai összefüggés állítható fel a h-index (h), a publikációk száma (n) és a publikációk átlagos idézettsége (x) között [Glänzel, 2006]:

$$h \approx n^{1/3} x^{2/3}$$

A h-index sajátossága, hogy egyetlen számmal jellemzi a publikációs termelékenységet ( $n$ ) és az idézettségi hatást ( $x$ ). Ez a sajátossága bizonyos értékelési feladatok esetén más mutatószámoknál alkalmasabbá, más feladatoknál alkalmatlanabbá teszi. Mint mindig, az elemzőnek egyedi döntést kell hozni, hogy egy feladatra a h-index, annak valamilyen változata, vagy valamilyen egészen más mutatószám a legalkalmasabb.

#### KITÉRŐ

Bár Hirsch a mutatószámot kimondottan kutatók egyéni értékelésére vezette be, a h-index mára messze túlnőtt szerzője eredeti szándékain. A „h-típusú” indexek, vagyis a h-indexszel azonos elven képezett mutatószámok alkalmazhatónak mutatkoztak nemcsak a tudománymetria más területein (folyóiratok, országok értékelése stb.), hanem a természettudományok, a közösségi média vagy éppenséggel a sport területén, sőt egészen általánosan, tetszőleges hálózatok jellemzésére is (lásd pl. [Schubert & Schubert, 2019]).

Visszaemlékezések szerint az asztrofizikus Sir Arthur Eddington az 1930-as években alkotott egy ilyen típusú mutatószámot kerékpározási teljesítményének mérésére. Azt számolta, hogy melyik az a legnagyobb  $N$  szám, ahány napon  $N$  mérföldnél többet tett meg.

#### ➤ Egyetemi rangsorok

Napjainkban a felsőoktatási intézmények számára nemcsak hiúsági kérdés az egymáshoz képesti rangsorokban elfoglalt helyezés. A legtöbb egyetem számára a hallgatói létszám alapvetően befolyásolja a bevételeket, és a hallgatók – legalábbis részben – a különféle kiadványokban megjelent rangsorok alapján választanak.

Ez a kereslet megteremtette a kínálatot. Számos olyan intézmény, fórum van, amelyik rendszeresen publikálja az egyetemi rangsorokat. A Wikipédia erről szóló oldala ([en.wikipedia.org/wiki /College\\_and\\_university\\_rankings](http://en.wikipedia.org/wiki/College_and_university_rankings)) 23 nemzetközi rangsort ismertet, és további 31 országban készítenek egy vagy több országos vagy regionális rangsort.

A három legjelentősebb rangsor a Quacquarelli Symonds (QS) rangsor, a Times Higher Education (THE) rangsor és a Academic Ranking of World Universities (ARWU) rangsor. Ezt a legutóbbit a Shanghai Jiao Tong Egyetemen készítik, és ezért Sanghaj-rangsornak is nevezik. Mind a három rangsorolás összetett (composite) mutatószámokon alapszik, vagyis számos különböző szempont alapján készített rangsorolást különböző súlyokkal véve átlagolnak. Mind a szempontok kiválasztása, mind a hozzájuk rendelt súlyok meglehetősen önkényesek és vitathatók, és nem is teljesen transzparenssek. Az biztos, hogy mind a háromban jelentős súllyal esnek latba a tudományometriai mutatószámok; közöttük nemcsak a publikációkon és idézeteken alapuló értékek, hanem pl. a tudományos fokozatok, díjak, vagy a nemzetközi kapcsolatokat jellemző mutatószámok is. A három nagy rangsorolás hiányosságait nem küszöböli ki, de valamelyest kiegyensúlyozza a University of New South-Wales (Sydney) által létrehozott Aggregate Ranking of Top Universities (ARTU), amely a korábban említett három fő rangsort kombinálja egy egyesített listává.

Az Európai Bizottság által 2014-ben elindított U-Multirank kezdeményezést ([www.umultirank.org](http://www.umultirank.org)) azért emelném ki, mert – ha kezdetlegesen is – megkísérel egy „személyre szabott” rangsort kialakítani: az összegyűjtött mutatószámokat a felhasználó igényei szerint súlyozva (mennyire fontos a magas színvonalú kutatás, az alacsony tandíj, a nemzetközi kapcsolatok stb.) alakítja ki a végső sorrendet.

## 5. Szakterületi sajátosságok a tudománymetriában

Össze lehet-e hasonlítani az almát a körtével? Mivel én személy szerint ki nem állhatom a körtét, tehát számomra az összehasonlítás triviális: a legrosszabb alma is jobb, mint a legjobb körte.

A lényeg talán az, hogy nem az elvont dolgokat kell összehasonlítani, hanem meghatározott minőségi vagy mennyiségi tulajdonságaikat. Aligha vitatja bárki, hogy össze lehet hasonlítani egy alma és egy körte színét vagy súlyát. A következő kérdés az összehasonlítás relevanciája. Milyen következtetést vonhatunk le az összehasonlítás eredményéből?

Ezeket a szempontokat kell szem előtt tartani a tudománymetriai összehasonlítások esetében is. Meg kell találni a mérhető és összemérhető tulajdonságokat, és gondosan meg kell fontolni az összehasonlítás eredményéből levonható következtetéseket.

A tudománymetriai mutatószámok szakterületek közötti összehasonlításának módszere általában az, hogy a mutatószámokat a saját szakterületükre vonatkozó sztenderdekhez kell mérni, és az így normalizált mutatószámokat lehet azután összemérni. Talán már nem kell újból hangsúlyozni, hogy a módszer megválasztásának alapelve az összehasonlítás céljának pontos ismerete kell, hogy legyen; e nélkül a leggondosabb méricskélés is öncélú és értelmetlen marad.

A szakterületi összehasonlítás előfeltétele az, hogy legyen a tudományterületeknek egy egységesen elfogadott, megbízható kategorizálása, és hogy a vizsgálandó elemeket (szerzők, publikációk, intézmények) egyértelműen hozzá lehessen rendelni ezekhez a szakterületi kategóriákhoz.

A tudományterületek klasszifikációjának számos rendszere ismert az információtudomány elméletében és a könyvtári gyakorlatban. A legáltalánosabb és legismertebb a Melvil Dewey nevéhez fűződő tizedes osztályozási rendszer, amelynek kidolgozásához a tudománymetria úttörőjeként bemutatott Samuel Bradford is hozzájárult. Érdekességként megemlíthető, hogy bár Dewey teljes neve (az író Jerome Klapka Jerome-hoz hasonlóan) a magyar szabadságharc hőséneke nevével (Melville Louis Kossuth "Melvil" Dewey), túlságosan büszkék nem lehetünk rá, mert tudományos eredményeit beárnyékolta a szexuális zaklatás, rasszizmus és antiszemitizmus vádja. Ezek a vádak sajnos többnyire igaznak bizonyultak. A Dewey-féle tizedes osztályozás egyetemes rendszere mellett (melynek Európában elterjedt változata az ETO) számos tudományterület dolgozta ki a saját osztályozási rendszerét; a közgazdaságtanban a *Journal of Economic Literature* által kezdeményezett JEL kódok, a fizikában 1970-től az American Institute of Physics által bevezetett PACS (Physics and Astronomy Classification Scheme) volt az irányadó, 2016-tól az American Physical Society által javasolt PhySH (Physics Subject Headings) rendszer lépett a helyébe; a matematikának és a számítástudománynak is megvan a saját besorolási rendszere (MSC, CCS).

A nagy szakirodalmi adatbázisok (Web of Science, Scopus) létrehozták saját szakterületi kategóriáikat; ezeket használják tudománymetriai szolgáltatásaikban is (Journal Citation Reports, Essential Science Indicators, illetve SCImago, CiteScore). A sokféle rendszer előnye, hogy egy-egy elemzési feladathoz meg lehet keresni a legalkalmasabb kategorizálási sémát; nagy hátránya viszont, hogy a különféle klasszifikációkkal készített elemzések többnyire egymással összehasonlíthatatlanok.

A tudománymetriai vizsgálatok céljaira olyan szakterületi kategóriák használata látszik a legcélszerűbbnek, amelyben

- a publikációk besorolása viszonylag egyszerűen megoldható;
- az egyes kategóriák összemérhető nagyságúak;
- minden kategóriába lehetőleg statisztikailag értékelhető mennyiségű publikáció kerül;
- az egyes kategóriákon belül a tudománymetria szempontjából releváns sajátosságok (pl. publikációs, idézettségi szokások) hasonlóak.

Az említett elvek érvényesítésével 2003-ban létrehoztunk egy 15 fő- és 67 alterületet tartalmazó rendszert. A 15 főterület közül 12 a természettudományi terület, 2 társadalomtudományi és 1 a művészeti és humán tudományoké [Glänzel & Schubert, 2003]. A cikket idéző közel 100 munka többsége hasznosnak találta a felosztást, a tényleges használathoz szükséges információk azonban keveseknek állnak csak rendelkezésre.

A könnyebben elérhető kommerciális vagy szabad hozzáférésű adatbázisok osztályozási rendszerei talán kevésbé felelnek meg az elvi követelményeknek, de hozzáférhetőségük pusztán ténye elegendő ahhoz, hogy az elemzések túlnyomó többségében ezek valamelyikét használják.

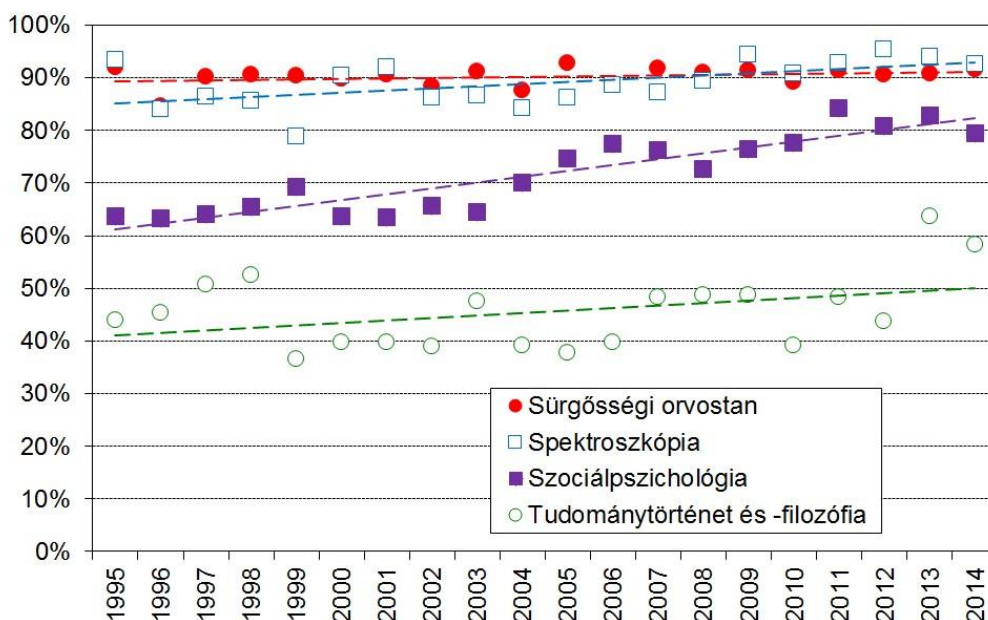
Melyek a leglényegesebb különbségek a tudománymetriai vizsgálatok szempontjából az egyes tudományterületek között?

- A szerzőket illetően a legfontosabb eltérések a szerzők számában, illetve – több szerző esetén – a sorrendjükben vannak. Egyes területeken (pl. a humán tudományokban vagy a matematikában) a cikkek számottevő része egyszerű szerzős; a másik véglet a már említett részecskefizika vagy a molekuláris genetika, ahol nem ritka a „megaszerzőség” (ezer feletti szerzőszámmal). A szerzői sorrend a matematikában és a fizika egyes (főként elméleti) területein rendszerint alfabetikus, az orvosi kutatásban a szerzők sorában elfoglalt hely és a kutatásban végzett szerep között hagyományos kapcsolat van. Más



területeken a szerzők száma és sorrendje nagy változatosságot mutat. A szerzők számának és sorrendjének a szerzői produktivitás, illetve a kutatókra szétosztott (frakcionált) teljesítmény-mutatók meghatározásában van fontos szerepe.

➤ A publikációk helye és típusa szerint főként a folyóiratokban és azokon kívül megjelent publikációk között kell különbséget tenni, illetve a folyóiratbeli publikációkon belül megkülönböztethetjük a különféle dokumentumtípusokat is. Általános tendencia, hogy a folyóiratcikkek részaránya a szakirodalomban folyamatosan növekszik. A szakterületi különbségek azonban nagyok (26. ábra).



26. ábra A folyóiratcikkek aránya négy kiválasztott szakterületen az 1995–2014 időszakban

[Schubert & Soós, 2015]

Látható, hogy a „kemény” természettudományos területeken ez a részarány végig 90% körül van, míg a kiválasztott humán területen a vizsgált időszak végére éri el az 50%-ot. A társadalom- és humán tudományok területén a nem

folyóiratban megjelent publikációk többsége könyv vagy könyvfejezet. Más területeken (főként a műszaki tudományokban és a számítástudományban) a konferenciakiadványok szerepe számottevő. Arról sem szabad elfeledkezni, hogy vannak olyan területek, ahol a kutatási eredmények megjelenési formája speciális: például a földrajztudományban a térképek, vagy a zenetudományban a kották.

A folyóiratokban megjelent cikkek sokszor speciális dokumentumtípusokba sorolhatók: pl. összefoglaló cikk (Review), levél (Letter) vagy szerkesztőségi cikk (Editorial). A szakterületek, sőt olykor az egyes folyóiratok hagyományai szabják meg, hogy ezek a besorolások pontosan mit fednek. Az összefoglaló cikkek általában új kutatási eredményeket nem tartalmaznak, de nagyon értékesek, mert az egyes kutatási témákban megjelent eredményeket összegyűjtik és értékelik. Hivatkozásjegyzékük általában hosszú (néha többszáz tételből áll), és többnyire gyakran idézik őket. Egyes folyóiratok (pl. az Annual Review... címűek) kizárólag ilyen cikkek közzlésére specializálódtak, de a folyóiratcím megtévesztő is lehet, mert pl. a nagyon jelentős Physical Review túlnyomóan eredeti kutatási eredményeket közöl. A folyóiratok által közölt levelek egy része valójában minimális tudományos információt közöl (afféle „kommentnek” tekinthető), de például a legtekintélyesebb orvosi folyóiratokba új tudományos eredmények rövid összefoglalását ilyen formában lehet beküldeni. Hasonlóképpen a szerkesztőségi közlemények is tartalmazhatnak formális bejelentéseket, de érdekes, értékes „vezércikk” tanulmányokat is.

➤ A szakterületi különbségek a hivatkozásokra/idézetekre is kiterjednek. A hivatkozásjegyzékek nyilvánvalóan rövidebbek a szűk kutatási területeken, amelyekkel esetleg csak néhány tucat kutató foglalkozik, mint ahol sokezer. Ennek megfelelően itt a kapott idézetek száma is lényegesen kevesebb. Egy másik szempont, ahol nagy különbségeket találunk, a hivatkozások kora.

Derek de Solla Price véleménye szerint az utolsó öt évben megjelent publikációkra vonatkozó hivatkozások aránya jól tükrözi egy tudomány „keménységét”. Bár részleteiben ez az állítás vitatható, de alapvető igazsága elfogadható. A publikációtípusokra vonatkozó szakterületi sajátosságok természetesen az idézett publikációkban is jól tükröződnek (27. ábra). A bal oldalon egy fizikai, a jobb oldalon egy szociológiai cikk irodalomjegyzékének tételei láthatók. Sötétpiros színnel vannak kiemelve az öt évnél újabb tételek.

**Conductance of crossed carbon nanotubes**

Nakanishi T, Ando T  
**JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN**  
 70 (6): 1647-1658 JUN 2001  
*Cited References*

Postma HWC, 2000, **PHYS REV B**, V62, P10653  
 Igami M, 2000, **PHYSICA B**, V284, P1746  
 Ando T, 2000, **SEMICON SCI TECH**, V15, pR13, DOI □  
 Führer MS, 2000, **SCIENCE**, V288, P494  
 Maarouf AA, 2000, **PHYS REV B**, V61, P11156  
 Choi HJ, 2000, **PHYS REV LETT**, V84, P2917  
 Yoshioka H, 2000, **PHYS REV B**, V61, P7316  
 Nakanishi T, 2000, **PHYSICA E**, V6, P872  
 Igami M, 2000, **MDL CRYST LIQ CRYST**, V340, P719  
 Ando T, 1999, **J PHYS SOC JPN**, V68, P3994  
 Igami M, 1999, **J PHYS SOC JPN**, V68, P3146  
 Ando T, 1999, **MICROELECTRON ENG**, V47, P421  
 Yoshioka H, 1999, **SYNTHETIC MET**, V103, P2527  
 Igami M, 1999, **J PHYS SOC JPN**, V68, P716  
 Odintsov AA, 1999, **EUROPHYS LETT**, V45, P598  
 Nakanishi T, 1999, **J PHYS SOC JPN**, V68, P561  
 Kwon YK, 1998, **PHYS REV B**, V58, P16001  
 Hertel T, 1998, **PHYS REV B**, V58, P13870  
 Kwon YK, 1998, **PHYS REV B**, V58, P13314  
 Matsumura H, 1998, **J PHYS SOC JPN**, V67, P3542  
 Ando T, 1998, **J PHYS SOC JPN**, V67, P2857  
 Ando T, 1998, **J PHYS SOC JPN**, V67, P1704  
 Delaney P, 1998, **NATURE**, V391, P466  
 NAKANISHI T, 1998, **PHYSICA B**, V249, P136  
 Egger R, 1997, **PHYS REV LETT**, V79, P5082  
 Kane C, 1997, **PHYS REV LETT**, V79, P5086  
 Ando T, 1997, **J PHYS SOC JPN**, V66, P3558  
 Nakanishi T, 1997, **J PHYS SOC JPN**, V66, P2973  
 Tamura R, 1997, **PHYS REV B**, V55, P4991  
 Tamura R, 1997, **SOLID STATE COMMUN**, V101, P601  
 Chico L, 1996, **PHYS REV B**, V54, P2600  
 Aji K, 1996, **J PHYS SOC JPN**, V65, P505  
 Saito R, 1996, **PHYS REV B**, V53, P2044  
 MINTMIRE JW, 1995, **CARBON**, V33, P893  
 BETHUNE DS, 1993, **NATURE**, V363, P605  
 IJIMA S, 1993, **NATURE**, V363, P603  
 AJIKI H, 1993, **J PHYS SOC JPN**, V62, P1255  
 IJIMA S, 1991, **NATURE**, V354, P56  
 DRESSLHAUS MS, 1988, **SPRINGER SERIES MAT**, V5  
 MINTMIRE JW, 1983, **PHYS REV LETT**, V50, P101  
 MINTMIRE JW, 1983, **PHYS REV B**, V28, P3283  
 SPAIN IL, 1973, **CHEMISTRY PHYSICS CA**, V8, P1  
 SLATER JC, 1954, **PHYS REV**, V94, P1498  
 IGAMI M, UNPUB J PHYS SOC JPN

**Perceptions of social support availability and coping behaviors**

Ueno K, Adams RG  
**SOCIOLOGICAL QUARTERLY**  
 42 (3): 303-324 SUM 2001  
*Cited References*

Ingram KM, 1999, **AIDS CARE**, V11, P313  
 LIN N, 1999, **HDB STUDY MENTAL HLT**, P241  
 TURNER RJ, 1999, **HDB SOCIOLOGY MENTAL**, P301  
 COLLINS RL, 1998, **HIV SOCIAL INTERACTI**, P30  
 DERLEGA VJ, 1998, **HIV SOCIAL INTERACTI**  
 SEIDEL J, 1998, **ETHNOGRAPH V5 O USER**  
 UENO K, 1998, **THESIS U N CAROLINA**  
 Bass LA, 1997, **J SOC PERS RELAT**, V14, P123  
 KANIYASTY K, 1997, **HDB PERSONAL RELATIO**, P595  
 SEIGEL K, 1997, **HEALTH PSYCHOL**, V16, P230  
 NORRIS FH, 1996, **J PERS SOC PSYCHOL**, V71, P489  
 Sarason IG, 1995, **NATO ADV SCI INST SE**, V80, P179  
 BARNES MK, 1994, **COMMUNICATION SOCIAL**, P175  
 BURLISON BR, 1994, **COMMUNICATION SOCIAL**  
 HAYS RB, 1994, **AIDS CARE**, V6, P379  
 PEARLIN LJ, 1994, **PSYCHOSOC REHABIL**, V17, P51  
 SARASON IG, 1994, **COMMUNICATION SOCIAL**, P91  
 TURNER HA, 1993, **J HEALTH SOC BEHAV**, V34, P37  
 HAYS RB, 1992, **J CONSULT CLIN PSYCH**, V60, P463  
 KESSLER RC, 1992, **SER CLIN C**, P259  
 SEIGEL K, 1991, **J HLTH SOCIAL BEHAV**, V32, P17  
 WELLMAN B, 1990, **AM J SOCIOL**, V96, P558  
 HAYS RB, 1990, **J COMMUNITY PSYCHOL**, V18, P374  
 HAYS RB, 1990, **AM J COMMUN PSYCHOL**, V18, P743  
 LAKEY B, 1990, **J PERS SOC PSYCHOL**, V59, P337  
 DUNKELSCHETTER C, 1990, **SOCIAL SUPPORT INTER**, P267  
 ECKENRODE J, 1990, **PERS RELATIONSHIP**, P83  
 CROCKER J, 1989, **PSYCHOL REV**, V96, P608  
 WEITZ R, 1989, **J HEALTH SOC BEHAV**, V30, P270  
 MATON KI, 1989, **AM J COMMUN PSYCHOL**, V17, P203  
 KURDEK LA, 1988, **J PERS SOC PSYCHOL**, V54, P504  
 ADELMAN MB, 1987, **COMMUNICATING SOCIAL**, P126  
 ANTONUCCI TC, 1986, **J CONSULT CLIN PSYCH**, V54, P432  
 SARASON IG, 1986, **J PERS SOC PSYCHOL**, V50, P845  
 WETHINGTON E, 1986, **J HEALTH SOC BEHAV**, V27, P78  
 LIN N, 1986, **SOCIAL SUPPORT LIFE**, P173  
 TAJFEL H, 1986, **PSYCHOL INTERGROUP R**, P7  
 WHEATON B, 1985, **J HEALTH SOC BEHAV**, V26, P352  
 ANTONUCCI TC, 1985, **HDB AGING SOCIAL SCI**, P94  
 COHEN S, 1985, **PSYCHOL BULL**, V98, P310  
 TARDY CH, 1985, **AM J COMMUN PSYCHOL**, V13, P187  
 THOITS P, 1985, **SOCIAL SUPPORT THEOR**  
 COHEN S, 1984, **HDB PSYCHOL HLTH**, V4, P253  
 LAZARUS RS, 1984, **STRESS APPRAISAL COP**  
 ROCKKS, 1984, **J PERS SOC PSYCHOL**, V46, P1097  
 SHINN M, 1984, **J SOC ISSUES**, V40, P55  
 SARASON IG, 1983, **J PERS SOC PSYCHOL**, V44, P127  
 THOITS PA, 1982, **J HEALTH SOC BEHAV**, V23, P145  
 KAHN RL, 1980, **LIFE-SPAN DEV BEHAV**, V3, P253

27. ábra A hivatkozások publikációtípus és kor szerinti szerkezete eltérő tudományterületeken.

A közkeletű vélekedés szerint a tudománymetria alkalmazhatóságának „vízválasztója” a hagyományos természettudományok (fizika, kémia, biológia) mellett az orvostudományokat, műszaki tudományokat és a matematikát is magában foglaló általánosabb értelemben vett természettudományok („STEM-fields”, „sciences”) és a társadalom- és humán tudományok („SSH”, „social sciences & humanities”) között húzódó határvonal. Amint láthattuk, a szakterületi különbségek a nagy tudománycsoportokon belül is jelentősek lehetnek; az almát az almával sem mindig könnyű összehasonlítani (28. ábra).



28. ábra Almák összehasonlítása

Egy reprezentatív mintán (kb. 1500-1500 publikáció a két nagy területről) összehasonlítottuk néhány tudománymetriai mutatószám alakulását a természettudományokban (Sci) és a társadalomtudományokban (SSci) [Schubert, 2012]. A várható különbségek mellett számos hasonlóság is feltűnhetett.

A részletes eredmények közül néhány az 1. függelékben elhelyezett 29-33. ábrákon látható.

Az elemzés fő tanulságaiként megállapíthattuk, hogy a társadalomtudományok területén

- A WoS adatbázis még a természettudományoknál is USA-centrikusabb;
- Az egyszerzős cikkek viszonylagosan nagy száma csökkenti a társszerzőségi hálózatok jelentőségét;
- A nyelvi és kulturális barrierek gátolják a közös tudásbázis kialakulását;
- A bibliográfiai hivatkozások száma, sőt funkciója is lényegesen különbözik az egyes szakterületek között;
- Az idézettségi életgörbe elhúzódása hosszabb vizsgálati időszakot tesz célravezetővé.

Ezek a különbségek az SSci tudományometriai elemzését gátolják, de nem teszik lehetetlenné.

## **6. A tudománymetria legújabb fejleményei és perspektívái**

A legutóbbi évtizedekben a tudományos publikálás évszázados módszerei gyökeresen átalakultak. A papíralapú folyóiratokat egyre nagyobb mértékben egészítik ki vagy váltják fel az elektronikus kiadványok, a korábban csak egészen szűk szakmai körben terjesztett, szakértői vélemények („peer review”) nélküli preprint kiadványok a legszélesebb nyilvánosság számára is hozzáférhetővé lettek, széles körben elterjedtek a publikációkat, illetve a szerzőket egyértelműen azonosító egységes rendszerek (DOI, ill. ORCID).

A tudománymetria, amely főképpen a tudományos publikációk metaadatainak vagy teljes szövegének kvantitatív elemzésével kísérli meg jellemezni a tudományos kutatás szerkezetét és dinamikáját, több mint fél évszázados története során módszereinek nagy részét a "hagyományos", nyomtatott folyóirat központú publikációs modellre fejlesztette ki. Nyilvánvaló, hogy a tudományos publikálás rendszerének változásait a tudománymetriának is követnie kell.

A tudományos publikációs színtér egy XXI. századi fejleménye a tudományos kutatás társadalmi ellenőrzésének az igénye. Ezt az igényt lényegében az táplálja, hogy a tudományos kutatás világszerte jelentős költségvetési támogatást emészt fel, és az adófizetők tudni szeretnék, hogy pénzüik hogyan hasznosul. Ennek a jogossága és realitása részben vitatható, de létezését tudomásul kell venni, és adekvát válaszokat kell rá találni. Ebben pedig a tudománymetriának is szerepe van.

A tudományos publikálás rendszerének legújabb fejleményei nem minden esetben üdvösek. Vannak, amelyek egyértelműen hasznosnak, vannak, amelyek tévesnek bizonyultak. A legtöbb azonban még alakulóban van, és végső kimenete kétséges. Minősítő állásfoglalásokat csak nagyon körültekintő mérlegelés után lehet tenni.

### **6.1. A nyílt hozzáférésű („open access”) publikációk**

A publikációs szcéna legújabb fejleményeinek egyike a nyílt hozzáférésű („open access”, OA) publikálás elterjedése. A Budapest Open Access Initiative 2001-es deklarációja óta a mozgalom (mert joggal nevezhetjük annak) számtalan hívet és legalább annyi kételkedőt – részben ellenzőt – szerzett magának. Míg a hagyományos publikálási modell szerint a folyóiratok ingyenesen közlik az arra érdemes cikkeket, és az olvasók előfizetési díj (vagy alkalmankénti vásárlási díj) ellenében olvashatják őket, az új modell az olvasást ingyenessé tenné, viszont a szerzőknek „cikkfeldolgozási díjat” (Article Processing Charge, APC) kell fizetniük a közlésért.

#### KITÉRŐ

Ifjúkoromban Budapesten két apróhirdetési újság működött. A címeikre már nem emlékszem. Az egyikben ingyenesen lehetett hirdetni, és a lap az újságárusoknál pénzért volt kapható, a másikat ingyenes szórólapként terjesztették, de a hirdetőknak fizetniük kellett a hirdetés megjelenéséért. A két üzleti modell – amely nyilván a tudományos publikálás két szóbanforgó változatának a megfelelője – hosszú évekig jól működött egymás mellett.

A tudományos publikálás speciális vonása, hogy itt az „eladók” és a „vevők” ugyanannak a kutatóközösségnek a tagjai; sarkítva fogalmazva a tudományos folyóiratokat ugyanazok olvassák, akik írják. Ez a specialitás a nyílt hozzáférés mellett és ellen is szolgáltat érveket.

A nyílt hozzáférésű publikálás híveinek azt az idealista elképzelését, hogy a cikkek ingyenes hozzáférése a nagy kiadóvállalatok profitéhségét valamelyest korlátozni tudja, erősen kétségessé teszi az a tény, hogy a nyílt hozzáférésű folyóiratok nagy részét ugyanazok a kiadók adják ki (vagy vásárolják fel), mint

az előfizetéseket. A „(Jobb felől üt) nekem füttyöl, (Bal felől üt) s nekem füttyöl” igazsága örök érvényűnek látszik.

A tudománymetria számára a modell megválasztásának nincs különösebben nagy jelentősége. Sok vizsgálatot végeztek arra nézve, hogy a nyílt hozzáférés növeli-e a cikkek idézettségét, de egyértelmű következtetésre nem sikerült jutni.

## **6.2. Publikációs-idézettségi adatbázisok, repozitóriumok**

A legfontosabb publikációs-idézettségi adatbázisokat (Web of Science, Scopus stb.) már korábban áttekintettük. A repozitóriumok közül kettő érdemel különös figyelmet.

Az ArXiv (arxiv.org) elektronikus preprint (e-print) repozitóriumot Paul Ginsparg, a Los Alamos National Laboratory (LANL) fizikusa hozta létre 1991-ben; 2011 óta a Cornell Egyetem könyvtára működteti. Eredetileg a fizika területén archiválta a megjelenés előtti kéziratokat, de mára a matematika, a csillagászat, a számítástudomány, a kvantitatív biológia és a kvantitatív közgazdaságtan is a gyűjtőköréhez tartozik. Lehetőséget ad megjegyzések hozzáfűzésére és akár több javított változat közzétételére is. A kezdeti ellenállás után a legtöbb kiadó elfogadja, hogy az ArXivra feltöltött anyagok nem számítanak publikációnak, ezért változatlan formában beküldhetők közlésre. A folyóiratban való megjelenés után a bibliográfiai adatokat az ArXivban fel kell tüntetni.

A PubMed Central (PMC) ingyenes digitális repozitórium. A PMC kizárólag ingyenesen és nyilvánosan hozzáférhető orvosbiológiai és élettudományi folyóiratcikkeket tartalmaz. Az Egyesült Államok Nemzeti Egészségügyi Intézete (NIH) hozta létre 2000-ben, eredetileg azzal a céllal, hogy a hozzá tartozó intézmények publikációit nyilvánossá tegye. Mára közel 4000 folyóirat teszi itt szabadon hozzáférhetővé a cikkeit – esetenként a publikálás után



bizonyos idő elteltével. Az egyszerű dokumentumtároláson túlmenően a PMC-be kerülő tételek egységes indexelési és formázási folyamaton mennek keresztül, és egységes azonosítóval ellátva, strukturált XML fájlként kerülnek az adatbázisba. Az idézetek közül a PMC kizárólag a rendszeren belüli cikkekben találhatóakat tárolja. A PMC nem tévesztendő össze a PubMed-del, a Medline és a PMC közös kereső rendszerével.

- *Kutatói közösségi hálózatok*

A kutatói hálózatok létrehozásának fő színterei a múltban a konferenciák és egyéb tudományos rendezvények voltak. Közhelynek számított, hogy a konferenciák legfontosabb mozzanatai nem az előadások, hanem a kávészünetek alatti beszélgetések voltak. Az itt megalapozott kapcsolatok azután tudományos együttműködésben, társszerzőségben folytatódhattak.

Az elektronikus kommunikáció a kapcsolatok építését lényegesen megkönnyítette. A spontán "hálózosodás" (levelezőlisták, blogok, fórumok) mellett létrejöttek a kutatói közösségi hálózatok intézményes formái is.

A ResearchGate ([www.researchgate.net](http://www.researchgate.net)) és az Academia.edu ([www.academia.edu](http://www.academia.edu)) egyaránt a kéziratok, cikkek megosztását, hatásuk követését, együttműködő partnerek keresését igyekeznek elősegíteni. Bár alapvető szolgáltatásaikat ingyenesen nyújtják, mind a kettő üzleti vállalkozásként működik (ebből a szempontból az Academia.edu doménnév végződése megtévesztő). Mindkét oldalt erős kritika éri azért, mert a szerzők által feltöltött cikkek jelentős része (alighanem több mint a fele) sérti a kiadók szerzői jogait.

2017. szeptemberében az International Association of Scientific, Technical, and Medical Publishers ügyvédi felszólítást küldött a ResearchGate-nek a szerzői jogok megsértésének megszüntetésére. A cég részlegesen eleget tett a

felszólításnak, és a szerzői jogokat sértő cikkek közül sokat eltávolított a szerzőkkel. A kiadók "óvatos optimizmussal" nyilatkoztak a fejleményekről.

A ResearchGate által a regisztrált kutatókra kiszámított és közzétett RG-mutatóról (RG score) és az ezzel kapcsolatos fenntartásokról már korábban volt szó.

A Mendeley ([www.mendeley.com](http://www.mendeley.com)) kézirat- és hivatkozásszerkesztő programból fejlődött online adatmegosztást, együttműködést, archiválást támogató rendszerré. Az adatok, kéziratok megtekintésének, letöltésének nyilvántartásával a kutatási eredmények "népszerűségének", a kutatóközösségre gyakorolt hatásának a mérését is elősegíti.

### **6.3. Egyéb szereplők**

A publikációk és szerzők azonosító rendszereiről (DOI és ORCID) már esett szó.

A tudományos publikációs folyamat egy nagyon lényeges, de eddig elhanyagolt területét, a bírálói ("peer review") és szerkesztői tevékenység nyomon követését célozza meg a Publons ([publons.com](http://publons.com)). Ez a kutatóknak szánt ingyenes szolgáltatás, amelyben nyomon követhetik, visszaigazolhatják és bemutathatják a világ tudományos folyóirataiban végzett bírálói és szerkesztői tevékenységüket. Két fiatal új-zélandi kutató hozta létre 2012-ben, és máig kétszázezer kutató csatlakozott hozzá. 2017-ben a Clarivate Analytics (a Web of Science jelenlegi tulajdonosa) vásárolta meg. A Publons célkitűzése az, hogy "kordában tartsa" a bírálói tevékenységet, és ezáltal felgyorsítsa a tudományos kommunikáció folyamatait. Azt hirdeti, hogy ha a bírálói, szerkesztői tevékenység a tudományos kutatás mérhető aspektusaként jelenik meg, ez elősegíti ezeknek a tevékenységeknek a nagyobb elismertségét, és ezáltal megnöveli a részt vevő kutatók tudományos presztizsét. Ez fokozza a

kutatók motivációját a bírálati folyamatban való részvételre, így gyorsítva a publikációs folyamatot. A bírálók maguk dönthetik el, hogy hozzájárulnak-e a bírálat tartalmának a nyilvánossá tételéhez, bár ezt a folyóiratok kiadói esetenként felülbírállhatják. A Publons partneri kapcsolatban van a legtöbb nagyobb folyóiratkiadóval (az Elsevier kivételével, amely 2014-ben a bírálói számára egy saját elismerési platformot hozott létre), valamint a folyóiratkiadáshoz kapcsolódó szolgáltatásokkal, mint az ORCID és az altmetric.com.

A Publons regisztrált kutatók szerkesztői és bírálói tevékenysége alapján egyéni, intézményi és országos rangsorokat tesz közzé. Érdekességként megemlíjtük, hogy Magyarország a regisztrált kutatók száma szerint Dél-Afrika és Szerbia között a 39. helyet foglalja el.

Az elmúlt évek egyik legvitatottabb kezdeményezése a Sci-Hub weboldal. Kazahsztáni egyetemi hallgatóként hozta létre 2011-ben Alexandra Elbakyan informatikusnő. Ezen a weboldalon jelenleg 62 millió tudományos publikáció teljes szövege érhető el – a szerzői jogi törvényeket semmibe véve. A Sci-Hub megítélése ellentmondásos. A kutatói közösség egy része az információszabadság élharcosának tartja, mások (főként a kiadók) törvénytörő kalóznak. Saját bevallása szerint Elbakyant a kutatási eredményekhez való szabad hozzáférés idealisztikus-kommunisztikus eszményei motiválják.

2015-ben az Elsevier New York államban beperelte a Sci-Hub-ot a szerzői jogok megsértéséért. A bírósági döntés következtében a Sci-Hub elvesztette eredeti doménnevét (sci-hub.org), azóta változó doménneveken működik. Maga Elbakyan ismeretlen helyen tartózkodik. Az Elsevier 12 millió dollár "vérdíjat" tűzött ki a nyomravezetőnek.

Megítélésének előjelétől függetlenül a Sci-Hub létezése megkerülhetetlen tényező a tudományos publikáció világában. A *Nature* kiadójának értékelése Elbakyant 2016-ban a tudomány tíz legbefolyásosabb személye közé sorolta.

A publikációs színtér újabb szereplőinek főbb adatait a 2. függelék táblázatában foglaltam össze.

# Irodalomjegyzék

Archambault É, Vignola Gagné, É (2004) *The Use of Bibliometrics in the Social Sciences and Humanities*. Science-Metrix, Montreal

Ball R, Tunger D (2006) Bibliometric analysis - A new business area for information professionals in libraries? *Scientometrics* 66(3):561–577.  
doi.org/10.1007/s11192-006-0041-0

Björneborn L, Ingwersen P (2004) Toward a basic framework for webometrics. *Journal of the American Society for Information Science* 55(14):1216–1227. doi.org/10.1002/asi.20077

Bradford SC (1934) Sources of Information on Specific Subjects, *Engineering: An Illustrated Weekly Journal (London)*, 137(26 January): 85–86.

Bradford SC (1946) *Romance of Roses*. F. Muller, London.

Bradford, SC (1948) *Documentation*. Crosby Lockwood, London.

Daud A, Muhammad F, Dawood H, Dawood H (2015) Ranking cricket teams. *Information Processing & Management* 51(2):62–73.  
doi.org/10.1016/j.ipm.2014.10.010

Franceschet M, Colavizza G (2017) TimeRank: A dynamic approach to rate scholars using citations. *Journal of Informetrics* 11(4):1128–1141.  
doi.org/10.1016/j.joi.2017.09.003

Gantmacher F (1975): *Lectures in Analytical Mechanics*. Mir, Moscow, ISBN 978-0846405511

Garfield E (1955) Citation indexes for science. A new dimension in documentation through association of ideas. *Science* 122(3159):108–111

Garfield E (1972) Citation Analysis as a Tool in Journal Evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science* 178(4060):471–479. doi.org/10.1126/science.178.4060.471

Glänzel W (2006) On the h-index - A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics* 67(2):315–321. doi.org/10.1007/s11192-006-0102-4

Glänzel W, Schubert A (2003) A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics* 56(3):357–367. doi.org/10.1023/A:1022378804087

*Glottometrics* (2002) Special issue to honor G.K. Zipf. 3:1–150.

Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* 102(46):16569–16572. doi.org/10.1073/pnas.0507655102

Kamarás K, Makara BG, Soós S (2021) A „tudomanymetria.com” módszerének alkalmazhatósága pályázatok értékelésére. *Magyar Tudomány* 182(4): 437–448. DOI: 10.1556/2065.182.2021.4.1

Kessler MM (1963) Bibliographic coupling between scientific papers. *American Documentation* 14(1):10–25. doi.org/10.1002/asi.5090140103

Lehmann R, Wohlrabe K (2017) Who is the ‘Journal Grand Master’? A new ranking based on the Elo rating system. *Journal of Informetrics* 11(3):800–809. doi.org/10.1016/j.joi.2017.05.004

Lotka AJ (1925) *Elements of Physical Biology*. Williams & Wilkins, ISBN 978-0598812681; Reprint kiadás (1956) *Elements of Mathematical Biology*. Dover, 1956.

Lotka AJ (1926) The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* 16(12):317–323.

Lundberg A et al. (2014) Entropy in the list of authors of scientific papers. *Annals of Improbable Research*, 2014, 20(1), 15–17.

[www.improbable.com/airchives/paperair/volume20/v20i1/Entropy-Authors-AIR-20-1.pdf](http://www.improbable.com/airchives/paperair/volume20/v20i1/Entropy-Authors-AIR-20-1.pdf)

Nalimov VV, Mulcsenko ZM (1969). *Наукометрия, Изучение развития науки как информационного процесса*. Nauka, Moscow. Magyarul (1980) Tudománymetria. Akadémiai Kiadó, Budapest.

Otlet P (1934) *Traité de documentation*, Mundaneum, Palais Mondial, Bruxelles.

Prathap G (2009) Is there a place for a mock h-index? *Scientometrics* 84(1):153–165. doi.org/10.1007/s11192-009-0066-2

Price, DJ de Solla (1963) *Little Science, Big Science*. Columbia University Press, New York, ISBN 9780231885751; Magyarul (1979) *Kis tudomány – Nagy tudomány*. Akadémiai Kiadó, Budapest.

Price, DJ de Solla (1985) *Little Science, Big Science...and Beyond*. Columbia University Press, New York, ISBN 978-0231049566.

Pritchard A (1969) Statistical Bibliography or Bibliometrics? *Journal of Documentation* 25(4): 348–349.

Schubert A (2012) *A szakterületi különbségek jelentősége a tudománymetriai elemzésekben (különös tekintettel a társadalom- és humán tudományokra)*. Előadás a Semmelweis Egyetem könyvtárában. [www.researchgate.net/publication/356732106\\_SchubertA2012jun29](http://www.researchgate.net/publication/356732106_SchubertA2012jun29)

Schubert A (2014) A tudománymetria kezdetei Magyarországon – Szigorúan személyes szemelvények. *Könyv és Nevelés* 16(1):21–31. [folyoiratok.oh.gov.hu/konyv-es-neveles/a-tudomanymetria-kezdetei-magyarorszagon](http://folyoiratok.oh.gov.hu/konyv-es-neveles/a-tudomanymetria-kezdetei-magyarorszagon)

- Schubert A (2015a) Az impaktfaktor és akiknek nem kell. *Orvosi Hetilap* 156(26):1065–1069. doi.org/10.1556/650.2015.30212
- Schubert A (2015b) A webometriáról – a 2015. évi Derek John de Solla Price- emlékérem odaítélése alkalmából. *Orvosi Hetilap* 156(36):1472–1474. doi.org/10.1556/650.2015.30226
- Schubert A (2016) Igazságosan vagy testvériesen? Szerzői szerepek és társszerzői részesedések a tudományos publikálásban. *Orvosi Hetilap* 157(13):512–516. doi.org/10.1556/650.2016.30418
- Schubert A (2019) Friedrich Engels, the great-grandfather of scientometrics. *ISSI Newsletter* 15(4):64–67.
- Schubert A, Schubert G (2016) A tudománymetria törvényalkotó „úttörői”. *Orvosi Hetilap* 157(2):74–78. doi.org/10.1556/650.2015.30279
- Schubert A, Schubert G (2019) All along the h-index-related literature: A guided tour. In: Glänzel W, Moed HF, Schmoch U, Thelwall M (szerk.) *Springer Handbook of Science and Technology Indicators*. Springer Handbooks. Springer, doi.org/10.1007/978-3-030-02511-3\_12
- Schubert A, Soós S (2015) A folyóiratcikkek súlya a különféle tudományterületek szakirodalmában: az elmúlt két évtized trendjei. *Orvosi Hetilap* 156(24):985–987. doi.org/10.1556/650.2015.30142
- Secker J, Boden D, Price G (szerk.) (2007) *The Information Literacy Cookbook. Ingredients, Recipes and Tips for Success*, ISBN 978-1843342250
- Soós Sándor (2012) Tudománymetriáról korszerű felfogásban. *Innotéka*, 2012. október 5. www.innoteka.hu/cikk /tudomanymetriarol\_korszeru\_felfogasban.499.html



Szabó AT (1985) Alphonse de Candolle's early scientometrics (1883, 1885) with references to recent trends in the field (1978–1983). *Scientometrics* 8(1–2):13–33. doi.org/10.1007/bf02025219

Todeschini R, Baccini A (2016) *Handbook of Bibliometric Indicators. Quantitative Tools for Studying and Evaluating Research*. Wiley, ISBN: 978-3-527-33704-0

Truesdell CA (1966): *Six Lectures on Modern Natural Philosophy*. Springer-Verlag, Berlin–Heidelberg, doi.org/10.1007/978-3-662-29756-8

Vladutz G, Cook J, (1984) Bibliographic coupling and subject relatedness. In: *Challenges to an Information Society, Proceedings of the 47th ASIS Annual Meeting*, White Plains: Knowledge Industry Publications, pp. 204–207.

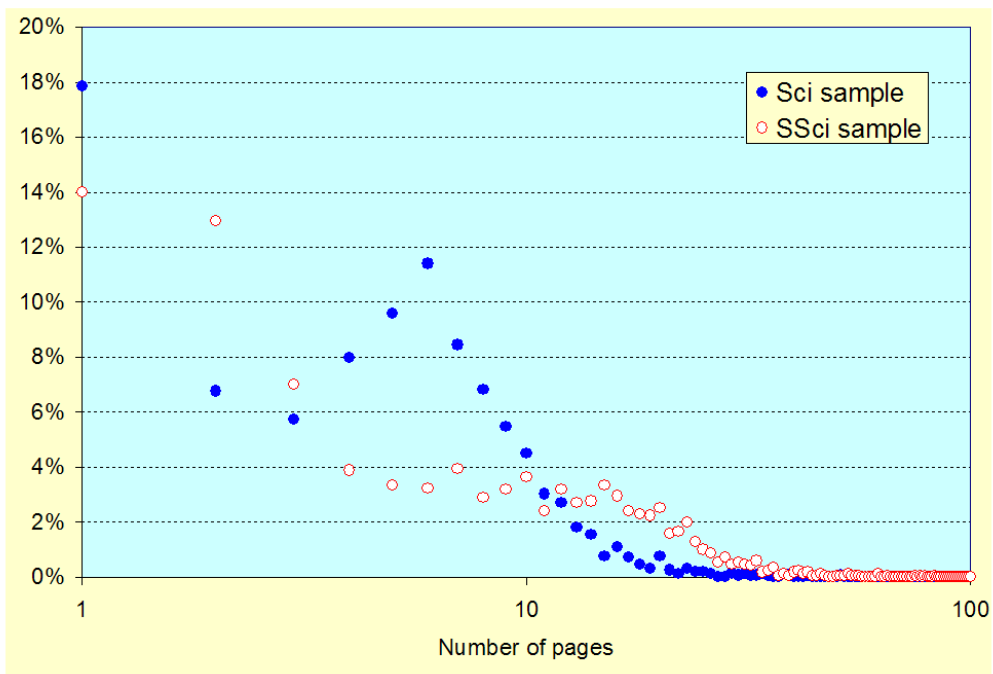
Waltman L, Noyons E (2018) *Bibliometrics for Research Management and Research Evaluation: A Brief Introduction*.  
www.cwts.nl/pdf/CWTS\_bibliometrics.pdf

Zipf GK (1935) *The psycho-biology of language*. Houghton Mifflin, Boston.

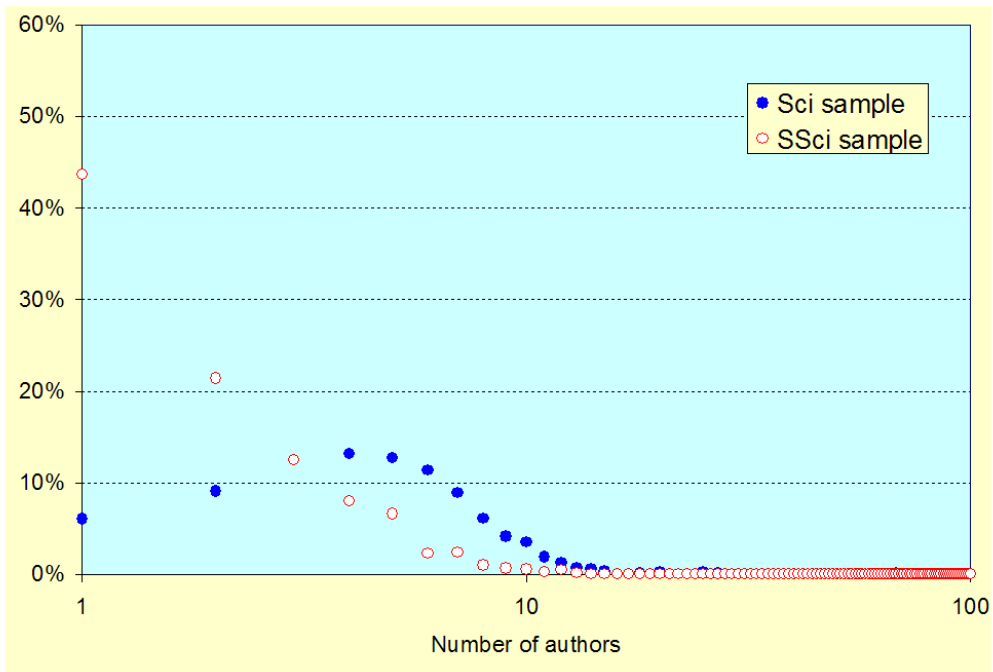
Zipf GK (1949) *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press, Cambridge.

# 1. függelék

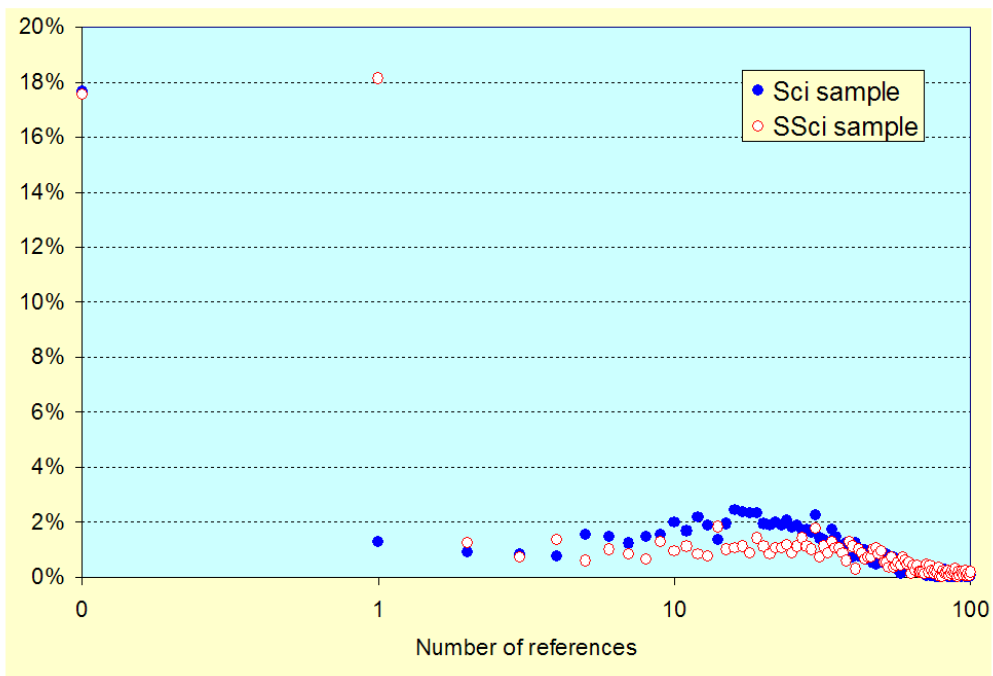
Az 1. függelék ábrái [Schubert, 2012] alapján készültek.



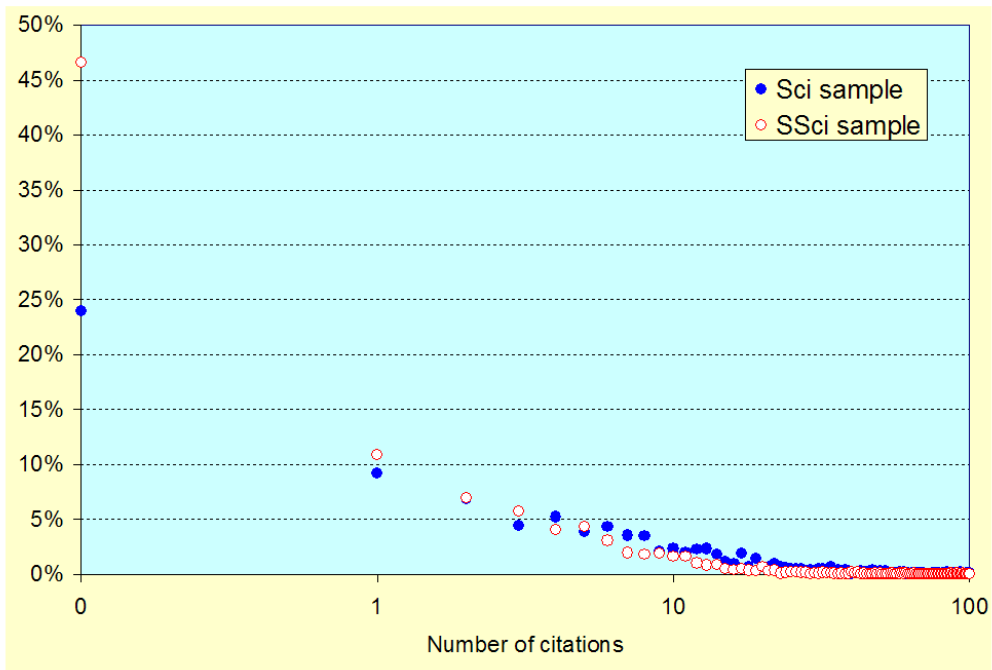
29. ábra A cikkek hosszúsága szerinti összehasonlítás



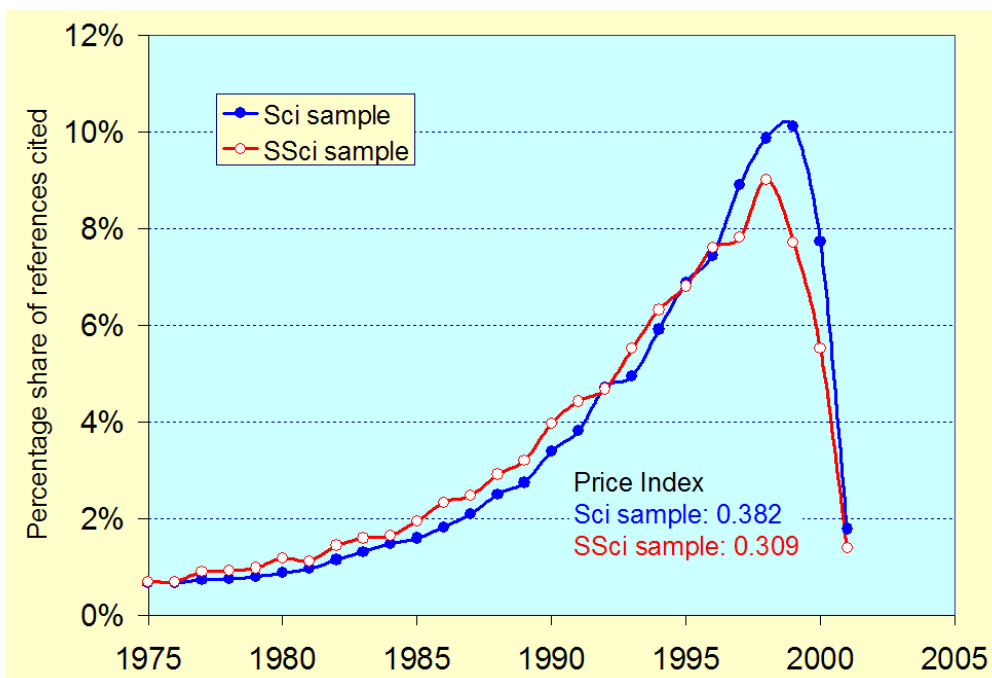
30. ábra A szerzők száma szerinti összehasonlítás



31. ábra A hivatkozások száma szerinti összehasonlítás



32. ábra A kapott idézetek száma szerinti összehasonlítás



33. ábra A hivatkozások kora szerinti összehasonlítás

## 2. függelék

A publikációs színtér újabb szereplőinek főbb adatai

	Tulajdonos / Működtető	Alapítás éve	Regisztrált felhasználók száma	Dokumentumok száma	URL
<b>Academia.edu</b>	Academia, Inc.	2008	150 millió	22 millió	www.academia.edu
<b>arXiv</b>	Cornell University	1991		2 millió	arxiv.org
<b>CiteSeer<sup>x</sup></b>	Pennsylvania State University	1997 CiteSeer; 2008 CiteSeerX		6 millió	citeseerx.ist.psu.edu
<b>Dimensions</b>	Digital Science	2018		100 millió	www.dimensions.ai
<b>DBLP</b>	Schloss Dagstuhl – Leibniz-Zentrum für Informatik	1993		5 millió	dblp.org/
<b>DOI (Digital Object Identifier)</b>	Nemzetközi DOI Alapítvány	1998		150 millió	www.doi.org
<b>Google Scholar</b>	Google, Inc.	2004		400 millió	scholar.google.com
<b>Mendeley</b>	Elsevier	2007	6 millió	40 millió	www.mendeley.com
<b>Meta</b>	Chan Zuckerberg Initiative	2010 (ScienceScape néven); 2017; 2022. márciusában megszűntetve		22 millió	www.meta.org

A publikációs színtér újabb szereplőinek főbb adatai (folytatás)

<b>Microsoft Academic</b>	Microsoft	2012 (Microsoft Academic Search); 2016; 2021. decemberben megszüntetve		220 millió	academic.microsoft.com
<b>ORCID (Open Researcher and Contributor Identifier)</b>	ORCID, Inc.	2012	10 millió		orcid.org
<b>Publish or Perish</b>	Anne-Wil Harzing	2007			harzing.com/resources /publish-or-perish
<b>Publons</b>	Clarivate Analytics	2012	200000	1 millió bírálat	publons.com
<b>PubMed Central (PMC)</b>	National Center for Biotechnology Information (NCBI)	2000		6,7 millió	www.ncbi.nlm.nih.gov/pmc/
<b>ResearchGate</b>	ResearchGate GmbH	2008	19 millió	130 millió	www.researchgate.net
<b>Sci-Hub</b>	Alexandra Elbakyan	2011		85 millió	sci-hub.se
<b>Scopus</b>	Elsevier	2004		75 millió	www.scopus.com
<b>Web of Science</b>	Clarivate Analytics	1964		171 millió	www.webofknowledge.com (csak regisztrált előfizetőknek)