

# *A tudományos tevékenység mérése*

– előadás –

Barta Zoltán

barta.zoltan@science.unideb.hu



zoology.unideb.hu

Általános kutatási ismeretek

PhD kurzus

Debreceni Egyetem

2020

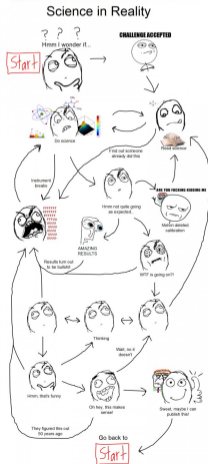
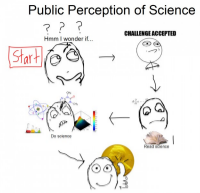
## VÁZLAT

- A tudományos tevékenység
- A tevékenység mérése
  - Miért?
  - Hogyan?
- A mérés következményei

# A tudományos tevékenység

# TUDOMÁNY

- a világ megismerése
- objektív (+/-)
- megismételhető (+/-)
- típusai:
  - felfedező tudomány
    - új ismeretek „termelése”
  - alkalmazott tudomány
    - ismeretek alkalmazása gyakorlati problémákra
- kollektív tevékenység

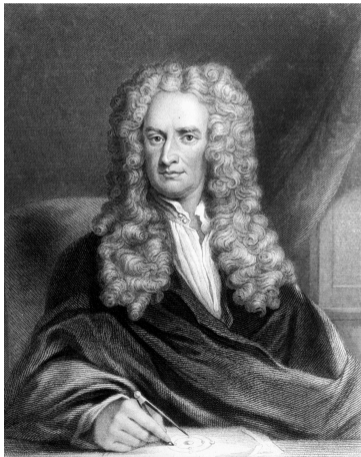


# KUTATÓK

*„If I have seen further, it is by standing on the shoulders of giants.”*

Sir Isaac Newton

- a tudomány „munkásai”
  - magasan képzettek
    - minimum MSc/MA,  
de legtöbbször PhD/DLA
    - specializált tudás
    - nyelvismeret
- együttműködnek egymással
  - korábbi eredmények



## PUBLIKÁCIÓK

- tudományos publikáció központi szerepe
  - „megtermelt” tudás közreadása
  - ellenőrzött keretek
    - *peer review*
  - eredmények ellenőrzésének lehetősége
    - objektivitás
    - ismételhetőség

### Publikáció a felfedező tudomány fő terméke!

- Nincs értelme kutatni publikáció nélkül!

*„the peer-reviewed research paper will remain a central research output that informs research assessment”*

# TUDOMÁNY FEJLŐDÉSE

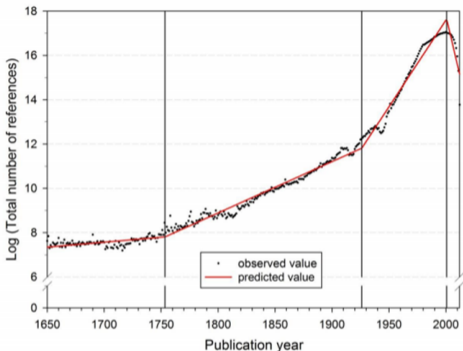


Figure 2. Segmented growth of the annual number of cited references from 1650 to 2012 (citing publications from 1980 to 2012)

- egyre gyorsuló ütem
- logaritmus skála az y tengelyen!

# A tudományos tevékenység mérése

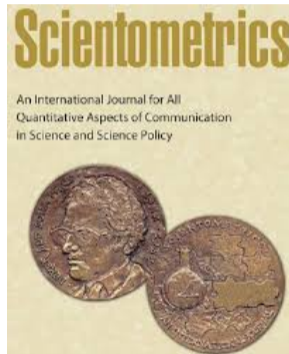
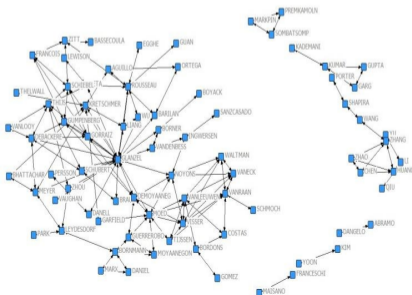






# MIÉRT KELL MÉRNI?

- tudomány társadalmi jelenséggé vált
  - tanulmányozható objektum
    - hogyan működik?
    - hogyan tehető hatékonyabbá?
  - külön tudományág: *scientometria*
    - saját folyóiratok, pl. *Scientometrics*



# SCIENTOMETRIA

## Történet

- Derek J. de Solla Price (1922 – 1983)
  - „the father of scientometrics”
- Eugene Garfield (1925 – 2017)
  - Science Citation Index

## Szakaszok

- publikációk analízise
- rangsorolás
  - egyetemek: ARWU, QS, THE
  - folyóiratok: IF
  - személyek: cikkek száma

## Módszerek

- indexek
- hálózatelemzés



# HOGYAN MÉRHEŐ A TUDOMÁNYOS TEVÉKENYSÉG?

- Fő cél: a tudományos *minőség* mérése
  - nehezen megragadható
    - eredetiség
    - új irányvonalak nyitása
    - szépség
  - csak *visszanézve* eldönthető
- **közelítések**
  - **produktivitás**
    - felfedező kutatások: *cikkek száma*
    - alkalmazott kutatások: *szabadalmak, termékek száma*
  - **hatás**
    - felfedező kutatások: *citációk*
    - alkalmazott kutatások: *profit*



# TUDOMÁNYMETRIAI SZINTEK

## Folyóirat-metria

- Journal Impact Factor (JIF)
- H-index
- SCimago Journal Rank (SJR)

## Cikk-metria

- idézettség
- szerzőszám

## Szerző-metria

- cikkek száma
- citációk száma
- h-index
- kumulatív impaktfaktor



# FOLYÓIRAT-METRIÁK

A három nagy játékos

- Journal Citation Report, Clarivate Analytics
  - <https://jcr.clarivate.com/>
  - fizetős, egyetemről elérhető
  - Journal Impact Factor (JIF)
- SCImago Journal Ranking, Scopus, Elsevier B. V.
  - <https://www.scimagojr.com/>
  - ingyenes
  - SCImago Journal Rank (SJR)
- Google Scholar, Google Ireland Limited
  - <https://scholar.google.com/>
  - ingyenes
  - H-index



Scopus Preview



# JOURNAL IMPACT FACTOR

- Animal Behaviour  
(2.674)

$$\text{JIF} = \frac{\text{Citations in 2018 to items published in 2016 (920) + 2017 (741)}}{\text{Number of citable items in 2016 (301) + 2017 (320)}} = \frac{1,661}{621}$$

apps.clarivate.com

„Experience has shown that in each specialty the **best journals** are those in which it is **most difficult** to have an article accepted, and these are the journals that have a **high impact factor**.”

Hoeffel, 1998

- magasabb JIF → nagyobb presztízs → nagyobb láthatóság
  - igyekezzünk területünk legmagasabb JIF-ú újságjaiba publikálni!
- DE ne felejtsük: JIF **manipulálható!**



# JIF KRITIKÁK

## JIF manipulálható!

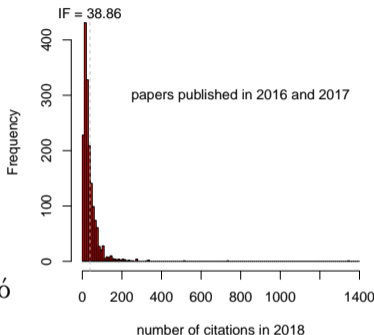
- szerkesztői összefoglaló cikk évente
- újságot idéző cikkek preferálása
- szerkesztők általi „zsarolás”: idézd ezt és ezt!
- cikkek évelejei tömeges publikálása

## Statisztikai tulajdonságok

- citációk: erősen ferde eloszlás
- átlag használata kérdéses
  - Nature 2018-s IF: 38.86
  - cikkek 68%-ánál kevesebb a citáció

## Összehasonlíthatóság

- csak azonos, sokszor szűk, területen belül
  - területek mérete eltér
  - citációs mintázatok eltérőek
- csak újságokat!



NATURE | NEWS

### Brazilian citation scheme outed

Thomson Reuters suspends journals from its rankings for 'citation stacking'.

Richard Van Noorden

27 August 2013

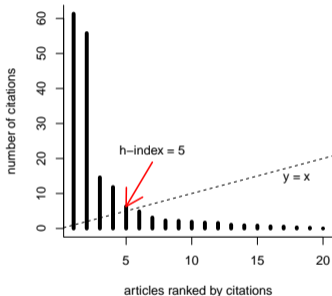
# H-INDEX

## Jellemzők

- **definíció:** legalább  $x$  olyan cikk van, amelyekre  $x$  vagy annál több citáció esik
- egyszerre méri a produktivitást és a hatást

## Kritikák

- manipulálható
  - mint az IF
- területek összehasonlíthatósága?
  - nem igazán vizsgált
- egész szám
  - csökkentett megkülönböztetés



# ÚJSÁGOK RANGSOROLÁSA

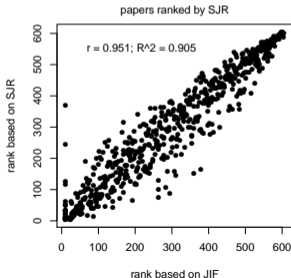
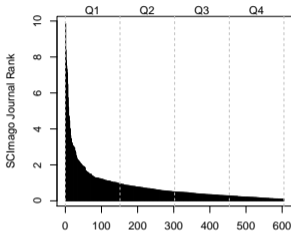
- újságok rangsorolása
  - csak adott tématerületen belül
  - valamilyen mérték alapján
    - SCImago Journal Rank, JIF
  - újságok kategóriákba sorolása
    - Q1: felső 25%; Q2: második 25%;
    - Q3: harmadik 25%; Q4: alsó 25%
    - esetleg D1: felső 10%

## Előnyök

- összevethető területek között

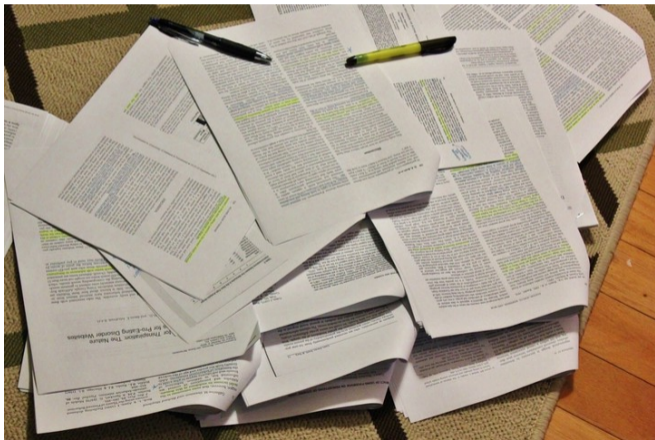
## Hátrányok

- nagyon durva skála
- kategória határon elenyésző a különbség



# *KÖZLEMÉNY SZINTŰ MÉRTÉKEK*

- idézettség
- szerzők száma



## ***IDÉZETTSÉG***

- hivatkozás jegyzék
  - az adott cikkben forrásként felhasznált művek jegyzéke
- típusai
  - függő hivatkozás
    - az idézett és idéző cikk szerzői között átfedés van
  - független hivatkozás
    - nincs átfedés az idézett és idéző cikk szerzői között

# HIVATKOZÁSTÍPUSOK

VOL. 155, NO. 4 THE AMERICAN NATURALIST APRIL 2000

## Daily Patterns of Optimal Producer and Scrounger Use under Predation Hazard: A State-Dependent

### Dynamic Game Analysis

### Literature Cited

Zoltán Barta<sup>1,\*</sup> and Luc-Alain Giraldeau<sup>2,†</sup>

független

függő

- elkülönítés nem mindig triviális
- Jeremy Smith vs John Smith  
→ J. Smith vs J. Smith

publikációs név megválasztása!

Baker, M. C., C. S. Belcher, L. C. Deutsch, G. L. Sherman, and D. B. Thompson. 1981. Foraging success in junco flocks and the effects of social hierarchy. *Animal Behaviour* 29:137–142.

Barnard, C. J., and R. M. Sibly. 1981. Producers and scroungers: a general model and its application to captive flocks of house sparrows. *Animal Behaviour* 29: 543–550.

Barnard, C. J., and D. B. A. Thompson. 1985. Gulls and plovers: the ecology of mixed species feeding groups. Croom Helm, London.

Barta, Z., and L.-A. Giraldeau. 1998. The effect of dominance hierarchy on the use of alternative foraging tactics: a phenotype-limited producing-scrounging game. *Behavioural Ecology and Sociobiology* 42:217–223.

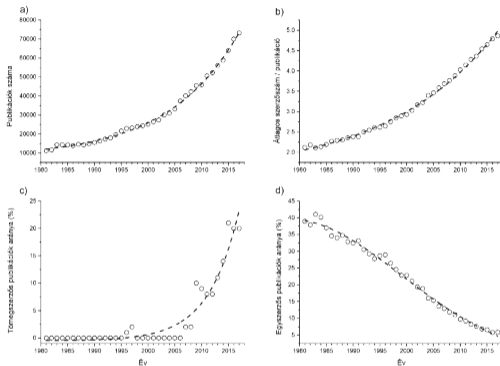
Caraco, T., and L.-A. Giraldeau. 1991. Social foraging: producing and scrounging in a stochastic environment. *Journal of Theoretical Biology* 153:559–583.

# IDÉZETTSÉG

## Problémák

- tudományterület-függő
  - kutatók száma, citációs szokások
- függ a cikk típusától
  - áttekintő cikkeket (*review*) jobban idézik
- függ szinte minden egyébttől (pl. cím hosszától)

# SZERZŐK SZÁMA



Magura 2019


- átlagos szerzőszám, tömegszerzős (50+) cikkek aránya rendületlenül nő
- egyszerűs cikkek aránya csökken



# SZERZŐ SZINTŰ INDEXEK

- cikkek száma
- hivatkozások száma
- h-index
- kumulatív impaktfaktor

Google Scholar SIGN IN



**Charles Robert Darwin** FOLLOW

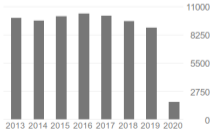
naturalist (1809-1882)  
Verified email at unr.edu.ar - [Homepage](#)  
[life sciences](#) [evolution](#) [biogeography](#) [speciation](#) [natural selection](#)

TITLE	CITED BY	YEAR
<a href="#">On the Orngin of Species</a> CR Darwin London: John Murray	52434*	1859
<a href="#">On rhe origin of species</a> CR Darwin Cambridge, MA: Harvard University Press	50648*	1859
<a href="#">77ie Origin of Species</a> C Darwin London: Murray	50518*	1859
	*	

[GET MY OWN PROFILE](#)

**Cited by** [VIEW ALL](#)

	All	Since 2015
Citations	194956	50781
h-index	119	65
i10-index	645	250



Year	Citations
2013	9000
2014	8500
2015	9000
2016	9500
2017	9500
2018	8500
2019	8000
2020	1500

## CIKKEK SZÁMA

- produktivitás mértéke
  - hatást nem jellemzi
- általában a kutató saját maga tartja nyilván
  - publikációs adatbázisok sokat segítenek
    - Google Scholar, Web of Knowledge, Scopus, MTMT
- erősen függ az életkortól, társszerzők számától
- nagy mértékben meghatározza a szerző teljesítményét

Mérték	$\rho$	$R^2$
összes hivatkozás	0.842	0.710
h-index	0.879	0.773
társszerzők száma	0.929	0.864
DI-s cikkek száma	0.739	0.546

Spearman  $\rho$  a cikkek száma és az adott mérték között,  $n = 23507$  szerző, MTMT.hu

# HIVATKOZÁSOK SZÁMA

- a kutató cikkeire kapott **összes** hivatkozás száma
  - a hatást jellemzi
- általában csak *független* hivatkozások
  - csak gondozott adatbázisoknál (pl. MTMT.hu) megbízható
  - független és függő hivatkozások sokszor erősen korrelálnak

## Manipulálható

- öncitációk magas száma
- citációs kartellek
  - egymás munkáinak idézése
- zsarolás bírálóként

Vaidyanathan, a computer scientist at the Vel Tech R&D Institute of Technology, a privately run institute, is an extreme example: he has received 94% of his citations from himself or his co-authors up to 2017, according to a [study in PLoS Biology this month](#)<sup>1</sup>. He is not alone.

Van Noorden & Singh Chawla, 2019

# HIVATKOZÁSOK SZÁMA

- jellemzően **adatbázisokból**
  - **Web of Knowledge, Scopus**
    - fizetős, de az egyetemről elérhető (Scopus: ingyenes limitált)
    - csak az indexelt folyóiratokbeli hivatkozások (pl. könyvekből nem)
  - **Google Scholar**
    - *ingyenesen* elérhető, népszerű
    - profil létrehozása után *automatizált*
    - csak elektronikusan elérhető dokumentumokból (nyomtatott források kiesnek)
    - viszonylag pontatlan, DE mindenkinél ugyanúgy tévedhet
- ugyanaz a szerző, különböző adatbázisok → eltérő számú hivatkozás
  - korrelálnak, DE Scholar magasabb
  - hivatkozások maximalizálása: mindet meg kell nézni

Google Scholar



Scopus Preview

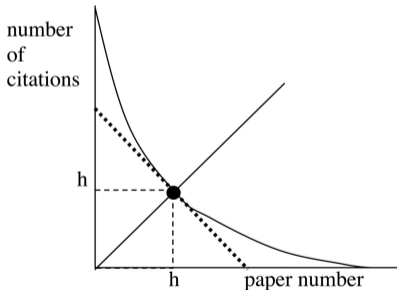
Clarivate  
Analytics

## H-INDEX

- kombinálja a produktivitást és a hatást
- Hirsch, J. E. (2005). "An index to quantify an individual's scientific research output". PNAS. 102 (46): 16569–16572.
- egy  $h$  indexű kutatónak van  $h$  db olyan cikke, amelyekre  $h$  vagy annál több hivatkozást kapott

Quantile	50%	75%	90%	95%	99%	100%
h-index	3	6	12	16	28	81

n = 23507, MTMT.hu



- több változat
  - összes vagy csak független citációk

## H-INDEX: KRITIKÁK

- citációk gyakoriságát nem veszi figyelembe
  - (mindkét h-index = 2)

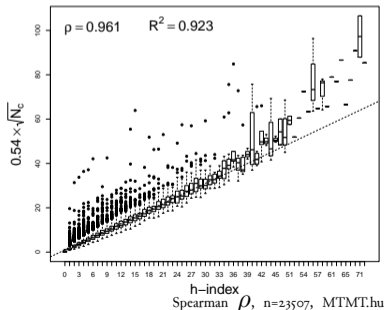
- erősen függ az összhivatkozástól

$$h \approx 0.54 \times \sqrt{N_c}$$

Yong, 2014

- területek összehasonlítására **nem alkalmas**
- **manipulálható!**
- határon lévő cikkek idézettsége

Cikkek	Szerző A	Szerző B
1	40	4
2	3	3
3	2	2
4	1	1
$\Sigma$	46	10



## *KUMULATÍV IMPAKTFAKTOR*

Egy  $n$  cikket írt kutató kumulatív impaktfaktora:

$$\sum_{i=1}^n F_i,$$

ahol  $F_i$  az  $i$ . cikket közlő újság impaktfaktora.

$F_i$ :

- vagy a cikk megjelenési évére számított IF,
- vagy a legfrissebb listából vett érték.

# KUMULATÍV IMPAKTFAKTOR

## Kritikák:

- területfüggő
  - *History* median IF: 0.500
  - *Cell & Tissue Engineering* median IF: 3.575
    - különbség ~ 7-szeres
  - **helyette**: újság rangok ( $Q_1$ ,  $Q_2$ , etc.), terület-függetlenek
- újság metrika alkalmazása személyre
  - DORA: San Francisco Declaration on Research Assessment, 2012



## Gondok:

- $Q_1$ ,  $Q_2$ , etc. rangok ugyanúgy újság metrikák alkalmazása személyekre
- magasabb IF-ű újságokban nehezebb publikálni
  - nagyobb a verseny: több kéziratot kapnak, jobban szelektálnak
  - IF mégis csak jelez valamilyen minőséget
    - DE: alacsonyabb IF-ű újságban is jelenhet meg kiváló cikk!!



## *SZERZŐ-METRIKÁK: ALTERNATÍVÁK*

- ne egyetlen metrikán alapuljon az értékelés!
- több szempont
  - cikkek száma
  - D<sub>I</sub>/Q<sub>I</sub> cikkek száma
  - citációk száma
  - első/utolsó szerzős cikkek
    - száma
    - ezekre kapott citációk száma
  - a szerző által öt legjelentősebbnek tartott cikk
    - citációi
    - rang
    - impaktfaktor
    - cikkek tartalma
- korrekció: tudományterületre, életkorra

A mérés következményei

# „PUBLISH OR PERISH”

- A tudomány személyes erőfeszítésből iparrá vált.
- Egy iparág működtetése ipari módszereket kíván.
  - ”vitality curve”, ”stack ranking”, ”rank and yank”
  - tudományban: ”publish or perish”
  - szereplők rangsorolása, felső 20-30% mindent visz
    - Pareto elv: kutatási eredmények 80%-át a kutatók 20%-a produkálja
- eredményes

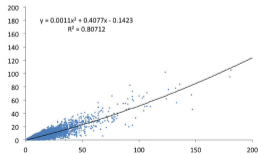


Fig. 8. Cited papers by total number of papers

Sandström & van den Besselaar, 2016

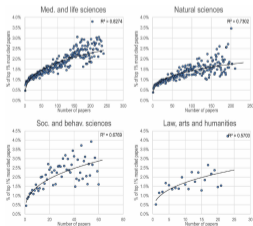


Fig. 9. Distribution of top 1% most cited papers (y axis), as a function of the number of papers published (x axis), for the subset of researchers who have produced their career number between 1981 and 1985, by domain. Only papers of researchers with 30 publications or more are shown. Power indices and  $R^2$  are here obtained using the Excel software.

doi:10.1016/j.joip.2016.02.002

Larivière & Costas, 2016

# HÁTULÜTŐK

hatalmas nyomás  
egyetlen értékmérő

} „felhívás keringőre”

- Goodhart's law

*„When a measure becomes a target, it ceases to be a good measure.”*

Marilyn Strathern

*„All metrics of scientific evaluation are bound to be abused.”*

Mario Biagioli, Nature, 2016

- „cobra effect”
- **törekvés a rendszer kijátszására**

*„It is unlikely that this unethical behavior can be totally eliminated because academics are a competitive, intelligent, and creative group of individuals.”*

Fong & Wilhite, PLoS ONE, e0187394

# *A RENDSZER KIJÁTSZÁSA*

## **Adathamisítás, plagizálás**

- több nagy botrány

## **Least publishable units**

- cikkek aprózása → számuk növelése
  - „szalámi cikkek”

## **Ki nem érdemelt szerzőség**

- honorárium szerzőség
- kieroszakolt szerzőség

## **Kartellek**

- több egyed összefog
  - érdemtelenül felveszik egymást társszerzőnek
  - egymás cikkeit hivatkozzák

# A RENDSZER KIJÁTSZÁSA

## Adathamisítás, plagizálás

- több nagy botrány

## Least publishable units

- cikkek aprózása → számok
- „szalámi cikkek”

## Ki nem érdemli

- honorárium
- kiegészítés

## Kartell

- több egy szerző
- érdemtelenség, társszerzőnek
- egymás cikkét

**Ne tedd ezt!!!**

# MELLÉKHATÁSOK

## A nyomás kihasználása

- predátor folyóiratok megjelenése
  - pénzért bármit publikálnak, pl. nincs rendes bíráló
  - [predatoryjournals.com](http://predatoryjournals.com), [bealllist.net](http://bealllist.net): lista a lehetséges predátorokról
- eltérített újságok
  - jó nevű, sokszor offline, újságok megjelenésének másolása
  - pénzért bármi publikálása
  - [predatoryjournals.com](http://predatoryjournals.com): lista a lehetséges eltérítésekről
- szerkesztői, cikkírási, könyvírasi felkérések özöne
  - sokszor ismeretlen kiadók
  - jó pénzért

## A jó oldal

- [retractionwatch.com](http://retractionwatch.com): visszavont cikkek követése
- [pubpeer.com](http://pubpeer.com): cikkek kritikája (J. Pruitt)

Köszönöm a figyelmet!



