



Kutatásértékelési szolgáltatások szakmai szemmel: tudománymetria kalauz az InCites megbízható felhasználásához

Soós Sándor

MTA KIK

Tudománypolitikai és Tudományelemzési Osztály

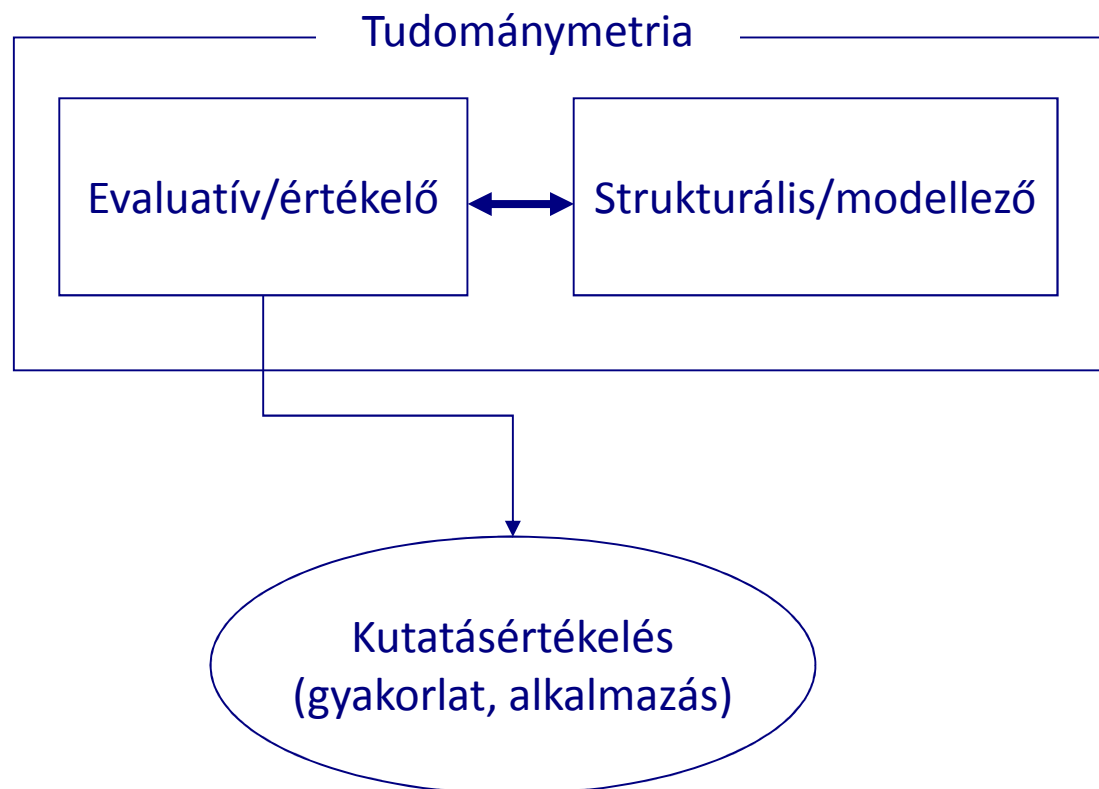


Tartalom

- A kutatásértékelési szolgáltatások kontextusa
- Mi az InCites?
- Elemzési szintek és egységek
- Mérőszámrendszer
- Mérés a kutatói, egyéni szinten
- Mérés az intézményi szinten: adatminőség
- Mérés az intézményi szinten: módszertan
- A jó gyakorlat



A kutatásértékelési szolgáltatások kontextusa





Egy informált gyakorlat tudománymetriaielemei

- **Cél:** sokszempontú értékelési profil szerkesztése a tudománymetria és tudománymodellezés eszközeivel
- **Felhasználás:** a döntéshozatal bemeneteként: [tudománymetriaileg] informált peer review”
- (Bibliometriaie) mutatórendszer dimenziói:
 - „*produktivitási*”, *kibocsátási mutatók*,
 - *Idézettségi hatás- és minőségmérők*
 - *strukturális mutatók* (kapcsolatrendszer, specializáció stb.)



Jellemző szintek és adminisztratív gyakorlat

	Értékelési szint (mikro, mezo, makro)		
Torzítások	Egyén (kutató)	Intézmény	Ország/Régió
Statisztikai megbízhatatlanság	H-index		
Kategóriahibák	IF/kumulált IF per		
Mechanisztikus alkalmazás	Nyers idézetszámok, H-indexek	Nyers idézetszámok, H-indexek	Nyers idézetszámok, H-indexek
Kontextuális tényezők elhagyása	Szakterületi profil	Szakterületi profil	Szakterületi profil
Közvetlen összemérés	Szakterületi profil	Szakterületi profil	Szakterületi profil
Adatminőség	Google Scholar	Vegyes adatok	Vegyes adatok
Alkalmas mérőszámok hiánya	X	X	X



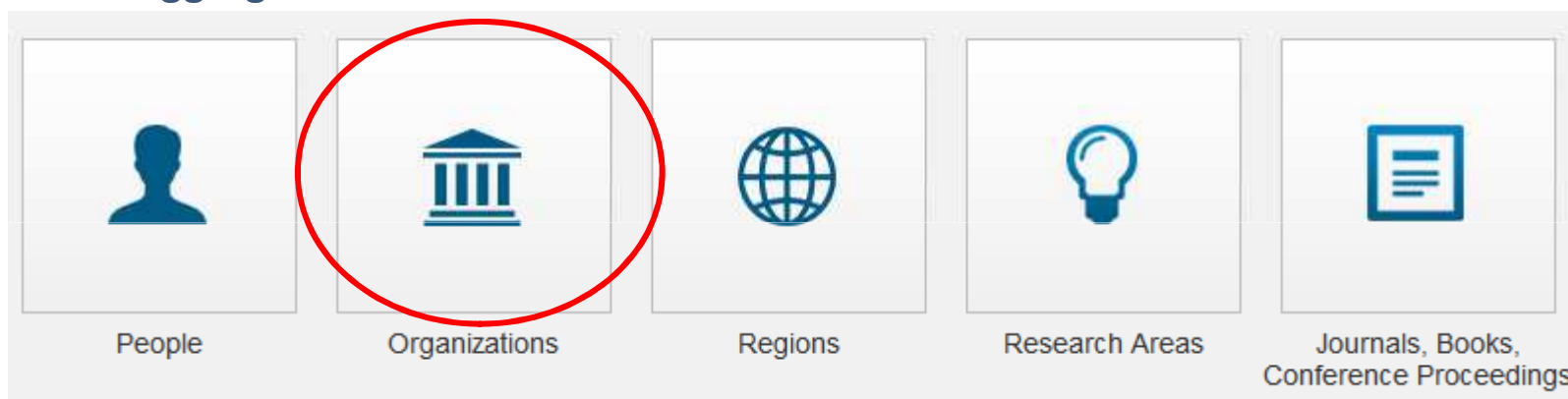
Mi az InCites?

- A Web of Science adattartalmára épülő platform, elsődlegesen kutatásértékelést támogató eszköztár
- **InCites:** a teljesítményértékelési dimenziók közül elsősorban a **hatásmérést** támogatja: mi, mikor, mit jelent?
- Kínálatában teret nyer a kutatómenedzsment további informálása: tudományos kapcsolatrendszer, kutatási profilt jellemző mutatók
- Alapvető működési séma: (1) Dokumentumhalmazok kiválasztása (WoS-ból) (2) ezek bibliometriai hatásmutatók széles körével való jellemzése és összemérése



Elemzési szintek és egységek: dokumentumhalmazok

- Aggregációs szintek



- Választható elemzési egységek: a WoS-közleményekben azonosítható
 - Szerzők
 - Affiliációkban megjelölt intézmények
 - Affiliációkban megjelölt „országok”
 - Lelőhelyek (vö. sikermutatók használata)



Mérőszámrendszer

- *Alapvető erény: szakmailag megalapozott, korszerű idézettségi mérőszámok*
- Idézettségi alapmutatók („alapstatisztikák”):
 - Times Cited; % Documents Cited; Citation Impact
- Normalizált hatásmutatók
 - Normalized Citation Impact; Journal Normalized Citation Impact; Average Percentile;
- Normalizált kiválósági mutatók
 - % Documents in Top 1%; % Documents in Top 10%; Average Percentile; % Highly Cited Papers; % Hot Papers
- Kombinált mutatók: H-index



Mérőszámrendszer

- *Strukturális mérőszámok*
- A tudományos kapcsolatrendszer mutatói:
 - International Collaborations; % International Collaborations;
% Industry Collaborations (!)
- Kutatási profil szerkezeti mutatói (specializáció)
 - Disciplinarity Index; Interdisciplinarity index





Mérés az egyéni/kutatói szinten

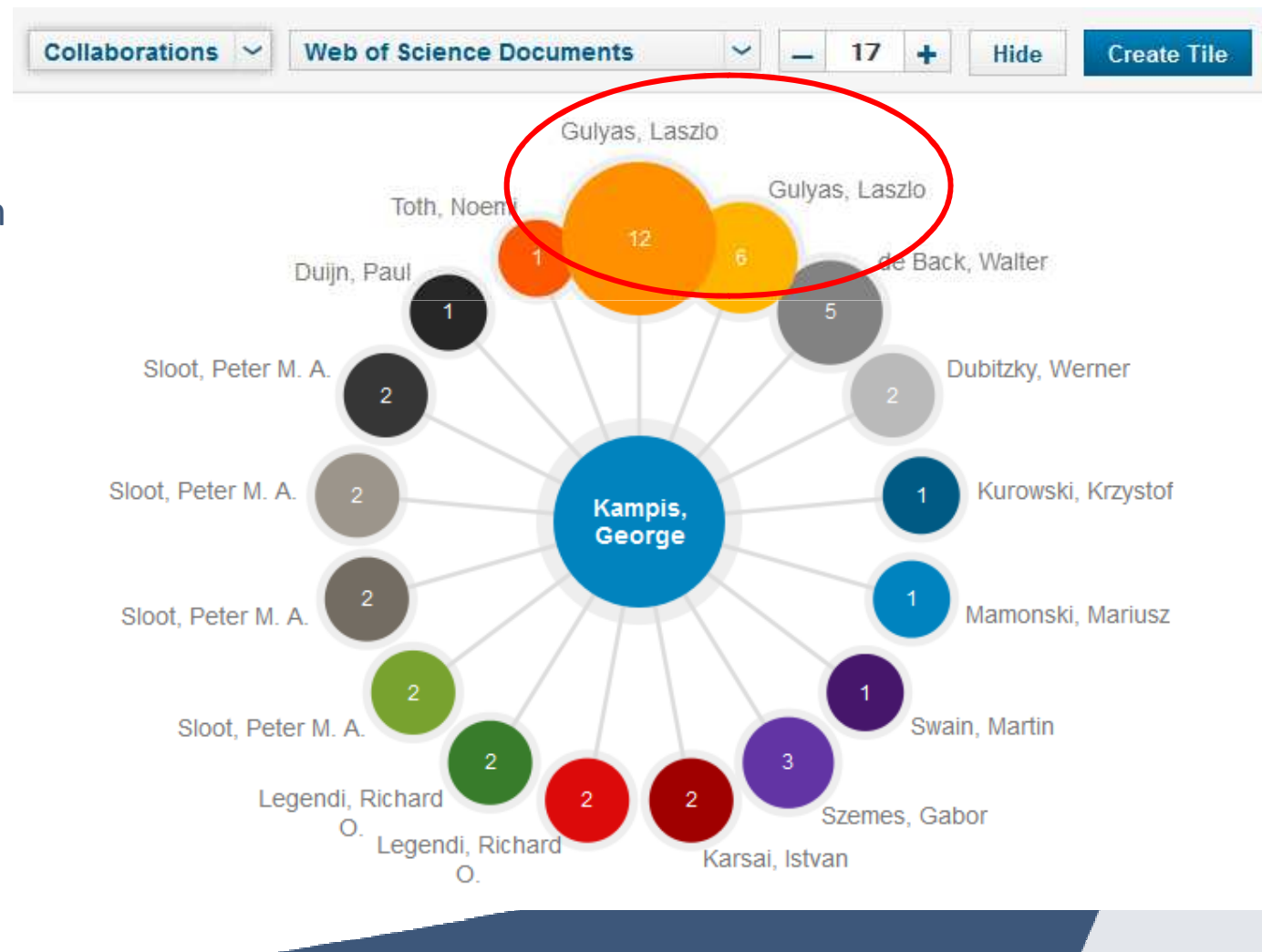
WoS:

Journal (paper-
based) Citation
Index

vs.

MTMT:

Author-based
citation index





Mérés az egyéni/kutatói szinten

- Az egyéni szintű tudománymetria: népszerű, szakmailag nehezen igazolható, jelenleg aktív kutatási terület

	Értékelési szint (mikro, mezo, makro)		
Torzítások	Egyén (kutató)	Intézmény	Ország/Régió
Statisztikai megbízhatatlanság	H-index		
Kategóriahibák	IF/kumulált IF per		
Mechanisztikus alkalmazás	Nyers idézetszámok, H-indexek	Nyers idézetszámok, H-indexek	Nyers idézetszámok, H-indexek
Kontextuális tényezők elhagyása	Szakterületi profil	Szakterületi profil	Szakterületi profil
Közvetlen összemérés	Szakterületi profil	Szakterületi profil	Szakterületi profil
Adatminőség	Google Scholar	Vegyes adatok	Vegyes adatok
Alkalmas mérőszámok hiánya	X	X	X



MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁR ÉS INFORMÁCIÓS KÖZPONT TUDOMÁNPOLITIKAI ÉS TUDOMÁNYELEMZÉSI OSZTÁLY

Mérés az intézményi szinten

	Web of Science Documents	Normalized Citation Impact	Times Cited	Docs Cited	Average Percentile	Inter- national Collabs	Docs in Top 10	Journal Normalized Citation Impact	Highly Cited Papers	Highly Cited Papers 1	Impact Relative to World	h index
Hungarian Academy of Sciences	● 26733	● 1,11	292036	73,36	57,07	14464	10,84	1,01	1,08	290	1,52	● 150
Semmelweis University	● 10475	● 1,03	95650	58,78	67,6	4688	8	0,95	0,76	80	1,27	● 108
Szeged University	● 9380	● 0,93	74714	64,96	64,37	4108	7,48	0,97	0,45	42	1,11	● 86
Eotvos Lorand University	● 9030	● 1,12	95100	69,1	59,52	4528	10	0,99	1,17	106	1,46	● 98
University of Debrecen	● 8856	● 1,02	80295	62,96	64,23	4168	8,9	0,97	0,99	88	1,26	● 91
Budapest University of Tech & Economics	● 8125	● 0,88	50913	61,51	63,69	2967	7,93	0,86	0,32	26	0,87	● 71
University of Pecs	● 5836	● 0,83	38112	55,28	70,88	2353	6,12	0,91	0,6	35	0,91	● 67
University of Pannonia	● 2441	● 1,11	16436	66,86	62,22	940	9,75	0,99	0,66	16	0,94	● 49
Szent Istvan University	● 2362	● 0,81	12169	61,9	67,66	855	6,77	0,93	0,17	4	0,72	● 41
Central European University	● 1439	● 1,07	5165	40,79	70,8	664	10,01	0,8	0,83	12	0,5	● 33
Corvinus University Budapest	● 1398	● 0,61	4802	49,14	73,04	432	4,72	0,79	0,14	2	0,48	● 27
University of Miskolc	● 929	● 1,05	2354	44,99	74,83	254	5,17	1,32	0,11	1	0,35	● 20
National Institute of Oncology Hungary	● 639	● 4,14	16696	59,15	63,85	289	15,49	1,49	3,6	23	3,63	● 48
Gedeon Richter Chemistry Works Limited	● 579	● 1,49	4087	65,8	66,27	130	8,12	1,59	0,52	3	0,98	● 28
University of West Hungary	● 554	● 0,69	2436	55,42	70,48	205	5,42	0,84	0,18	1	0,61	● 23
Kaposvar University	● 501	● 0,69	1852	55,29	74	189	3,39	0,89	0	0	0,51	● 20
Obuda University	● 315	● 1,93	253	27,62	81,36	83	7,62	1,06	0,32	1	0,11	● 7
University of Istvan Szechenyi	● 23	● 0,12	12	21,74	89,53	1	0	0,92	0	0	0,07	● 2



Mérés az intézményi szinten

ID	Measures	1	2	3	4	5	6	7	8	9	10	11	12
1	Web.of.Science.Documents	1,00	0,16	0,95	0,64	0,71	0,99	0,49	0,04	0,59	0,93	0,69	0,95
2	Normalized.Citation.Impact	0,16	1,00	0,27	0,37	0,46	0,14	0,80	0,78	0,70	0,37	0,48	0,28
3	Times.Cited	0,95	0,27	1,00	0,70	0,81	0,96	0,63	0,11	0,75	0,98	0,85	1,00
4	Documents.Cited	0,64	0,37	0,70	1,00	0,91	0,63	0,58	0,40	0,56	0,69	0,81	0,74
5	Average.Percentile	0,71	0,46	0,81	0,91	1,00	0,72	0,77	0,30	0,74	0,80	0,86	0,83
6	International.Collaborations	0,99	0,14	0,96	0,63	0,72	1,00	0,51	-0,01	0,63	0,93	0,71	0,96
7	Documents.in.Top.10	0,49	0,80	0,63	0,58	0,77	0,51	1,00	0,45	0,93	0,70	0,73	0,64
8	Journal.Normalized.Citation.Impact	0,04	0,78	0,11	0,40	0,30	-0,01	0,45	1,00	0,38	0,20	0,42	0,13
9	Highly.Cited.Papers	0,59	0,70	0,75	0,56	0,74	0,63	0,93	0,38	1,00	0,82	0,82	0,75
10	Highly.Cited.Papers.1	0,93	0,37	0,98	0,69	0,80	0,93	0,70	0,20	0,82	1,00	0,84	0,98
11	Impact.Relative.to.World	0,69	0,48	0,85	0,81	0,86	0,71	0,73	0,42	0,82	0,84	1,00	0,85
12	h.index	0,95	0,28	1,00	0,74	0,83	0,96	0,64	0,13	0,75	0,98	0,85	1,00

- A mutatók rangkorrelációja a hazai intézmények körében



Kitekintés: a normalizált hatásmutatók jelentése

- A tényleges hatás mérőszámcsaládja, a **hatás kontextusban mérése**: *szakterületre* (ill. a mű *korára és típusára*) normalizált idézettség
- **Háttér:**
 - Idézettség szakterületi normalizálása = a szakterületek összemérhetővé tétele
 - A legalapvetőbb értékelési torzítás (közlési viselkedésből): szakterületek eltérő idézéssűrűsége (vö. Matematika és klinikai orvostudomány)
 - Általános elv: az idézettség szakterületi normához viszonyítása:

$C_{norm}(P) = \text{mérhető Idézettség } (P) / \text{Várható idézettség } (C, Y, T)$
Y=P megjelenési éve, T=P típusa, C=P szakterületi kategóriája



Kitekintés: a normalizált hatásmutatók jelentése

- **Paraméteres változat: Szakterületi átlaghoz viszonyított idézetszám (Normalized Citation Impact)**
 - A hivatkozási eloszlás jellemző paraméterét (átlag) alkalmazza
 - Fő probléma: ferde eloszlások, az átlag torzító hatása
- **Nemparaméteres változat: Szakterületi eloszlásban elfoglalt pozíció**
 - Percentilis osztályok: A szakterület legidézettebb felső 1%-, 10%-ához való tartozás
 - Percentilis rang (Percentile rank)
 - Nem paraméteren alapszik

Kitekintés: a normalizált hatásmutatók jelentése

- Paraméteres normalizálás (Radicchi–Castellano, 2012, PLOS)

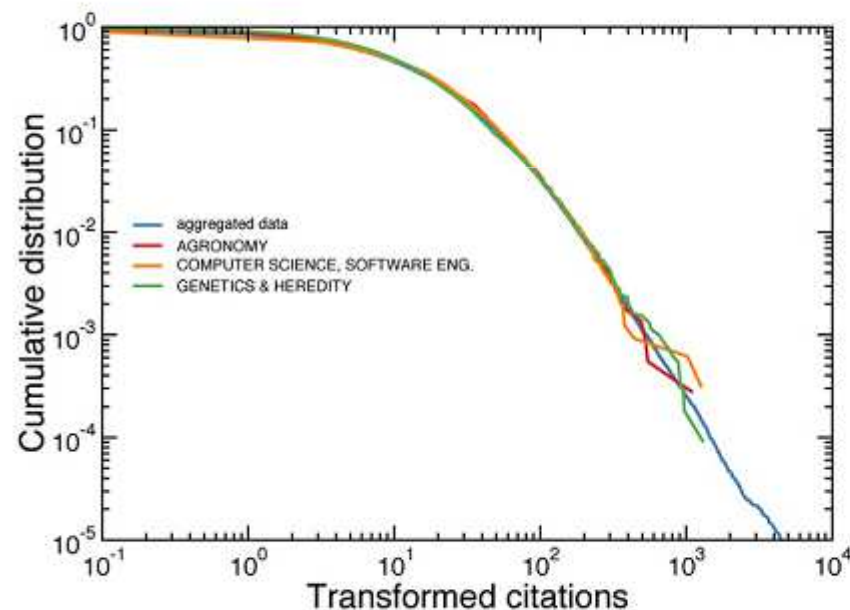
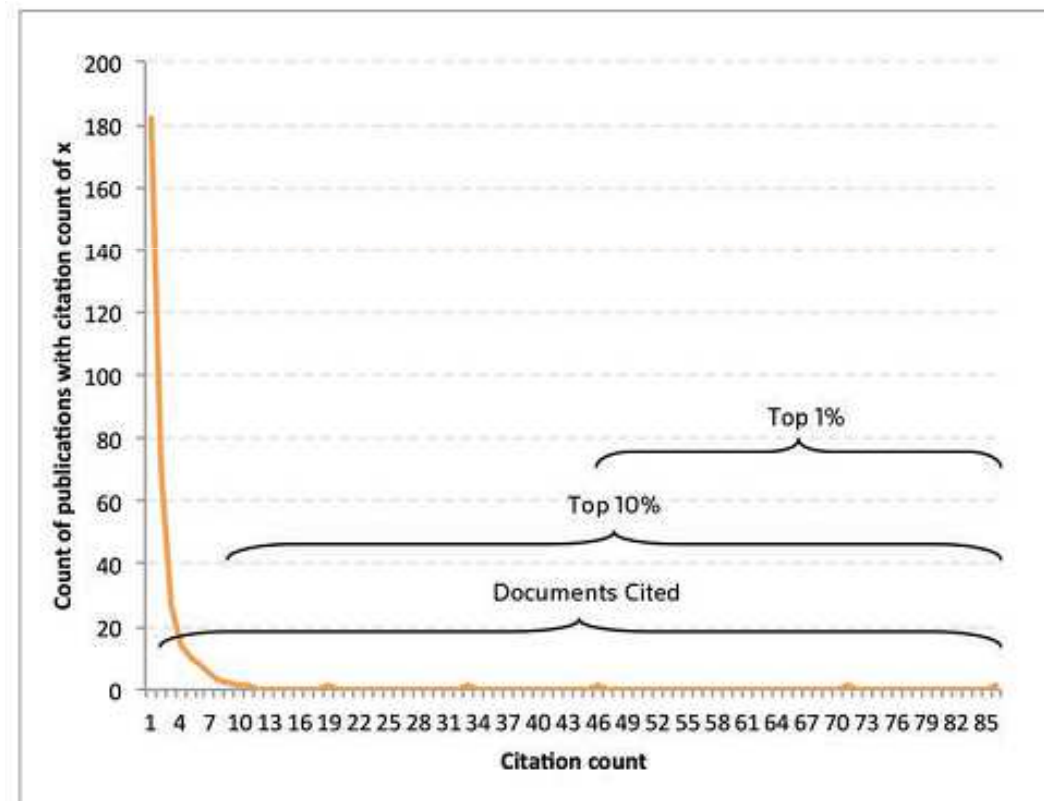


Figure 3. Cumulative distribution of the transformed citation counts. When raw citation numbers are transformed according to Eq. 2, the cumulative distributions of different subject-categories become very similar. All citation distributions are mapped on top of the cumulative distribution obtained by aggregating all subject-categories together (the common reference curve in the transformation). We consider here the same subject-categories as those considered in Figs. 1 and 2. The complete analysis of all subject-categories and years of publication is reported in the Supporting Information S2, S3, S4, S5, S6, and S7.
doi:10.1371/journal.pone.0033833.g003



Kitekintés: a normalizált hatásmutatók jelentése

- nemparaméteres normalizálás





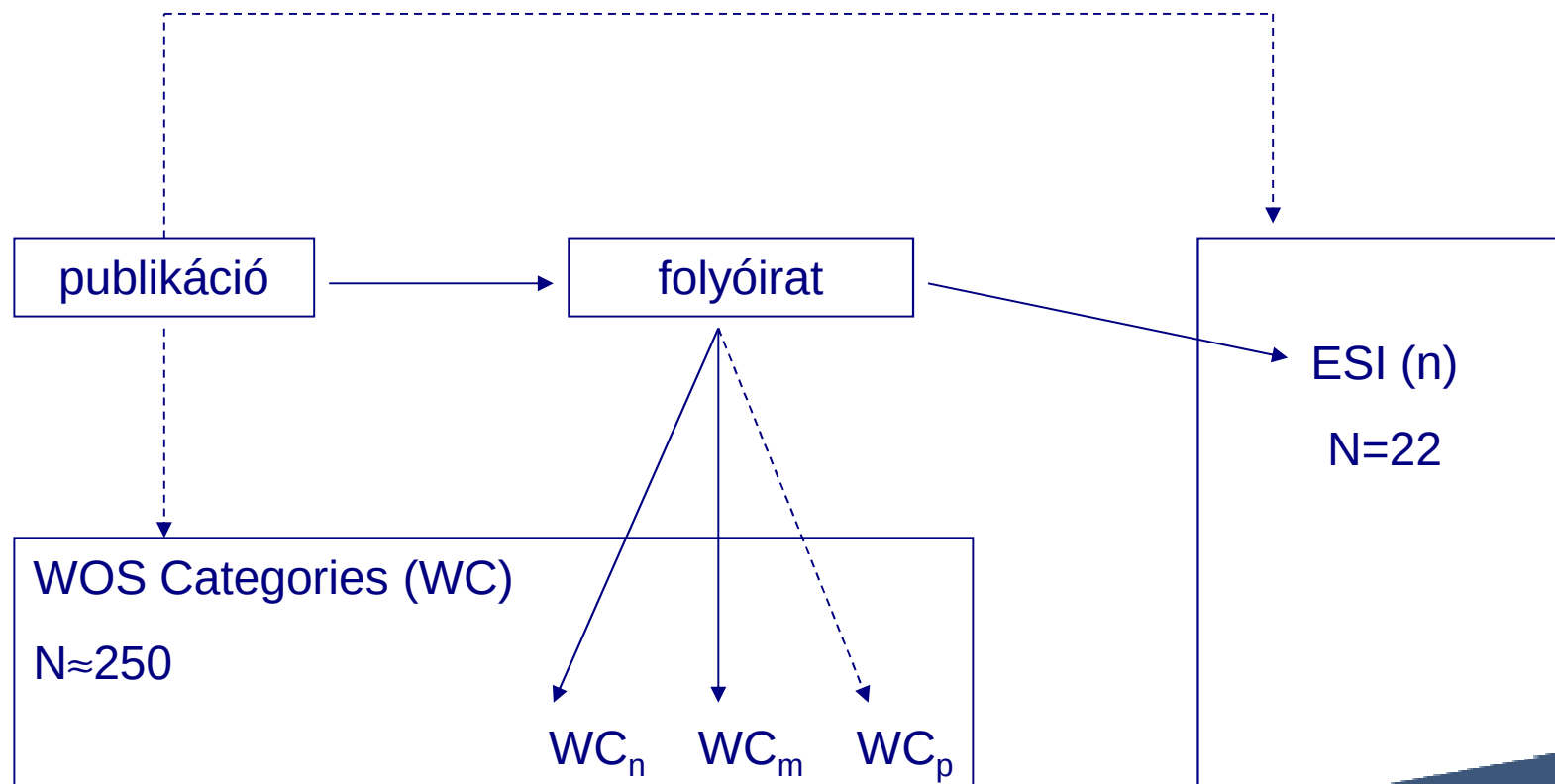
Kitekintés: a normalizált hatásmutatók jelentése

- **Mi az a szakterület (referenciahalmaz)? Elvileg:**
 - A szakterülethez (nemzetközileg) hozzárendelhető, idézési/hivatkozási, kibocsátási viselkedésében egységesként kezelhető, belső összemérhetőséget biztosító nemzetközi – ált. maximális – közleményhalmaz
- **Referenciahalmazok a gyakorlatban: folyóirat-kategóriák**
- **Kapcsolat szakterület és közlemény között: folyóirat**



Kitekintés: a normalizált hatásmutatók jelentése

- InCites-ban elérhető főbb szakterületi sémák





Kitekintés: a normalizált hatásmutatók jelentése

- **Összegzés: a normalizált hatásmutatók képzése**
 - Az intézményi (egyéni stb.) közleményhalmaz **heterogén** (szakterület, kor, típusok). Az idézettség fenti normalizálásai ezért az egyes közlemények szintjén zajlanak (item-oriented)
 - A közlemények normalizált idézettségi értékei intézményi (egyéni stb.) szintű aggregációban jellemzik a teljes közleményhalmazt (átlag, súlyozott átlag)
 - Eredmény: (intézményre) normalizált átlagos hatás, átlagos percentilis rang.



Mérés az intézményi szinten

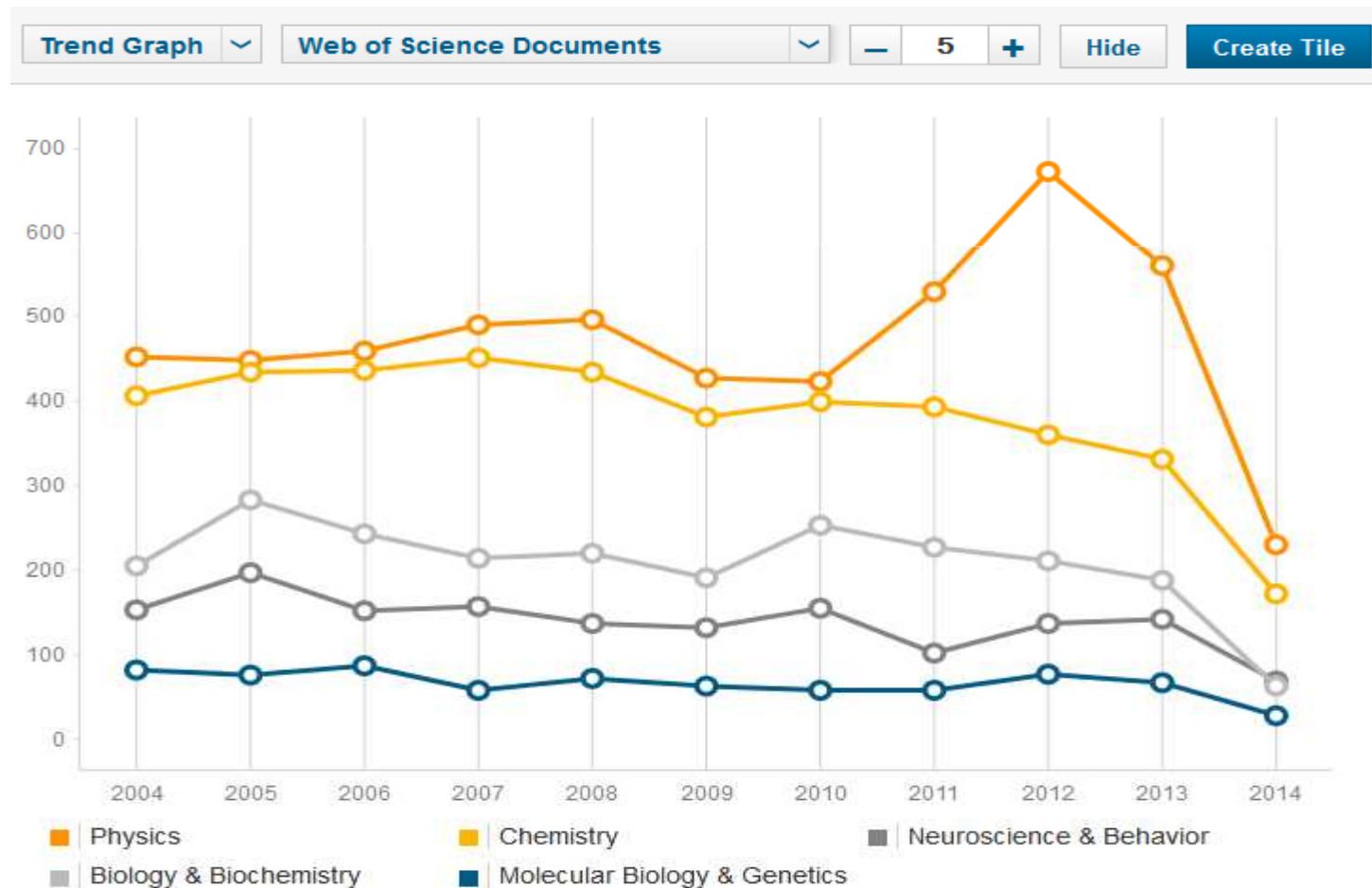
ID	Measures	1	2	3	4	5	6	7	8	9	10	11	12
1	Web.of.Science.Documents	1,00	0,16	0,95	0,64	0,71	0,99	0,49	0,04	0,59	0,93	0,69	0,95
2	Normalized.Citation.Impact	0,16	1,00	0,27	0,37	0,46	0,14	0,80	0,78	0,70	0,37	0,48	0,28
3	Times.Cited	0,95	0,27	1,00	0,70	0,81	0,96	0,63	0,11	0,75	0,98	0,85	1,00
4	Documents.Cited	0,64	0,37	0,70	1,00	0,91	0,63	0,58	0,40	0,56	0,69	0,81	0,74
5	Average.Percentile	0,71	0,46	0,81	0,91	1,00	0,72	0,77	0,30	0,74	0,80	0,86	0,83
6	International.Collaborations	0,99	0,14	0,96	0,63	0,72	1,00	0,51	-0,01	0,63	0,93	0,71	0,96
7	Documents.in.Top.10	0,49	0,80	0,63	0,58	0,77	0,51	1,00	0,45	0,93	0,70	0,73	0,64
8	Journal.Normalized.Citation.Impact	0,04	0,78	0,11	0,40	0,30	-0,01	0,45	1,00	0,38	0,20	0,42	0,13
9	Highly.Cited.Papers	0,59	0,70	0,75	0,56	0,74	0,63	0,93	0,38	1,00	0,82	0,82	0,75
10	Highly.Cited.Papers.1	0,93	0,37	0,98	0,69	0,80	0,93	0,70	0,20	0,82	1,00	0,84	0,98
11	Impact.Relative.to.World	0,69	0,48	0,85	0,81	0,86	0,71	0,73	0,42	0,82	0,84	1,00	0,85
12	h.index	0,95	0,28	1,00	0,74	0,83	0,96	0,64	0,13	0,75	0,98	0,85	1,00

- A mutatók rangkorrelációja a hazai intézmények körében



A mutatók érzékenysége a kiválasztott paraméterekre

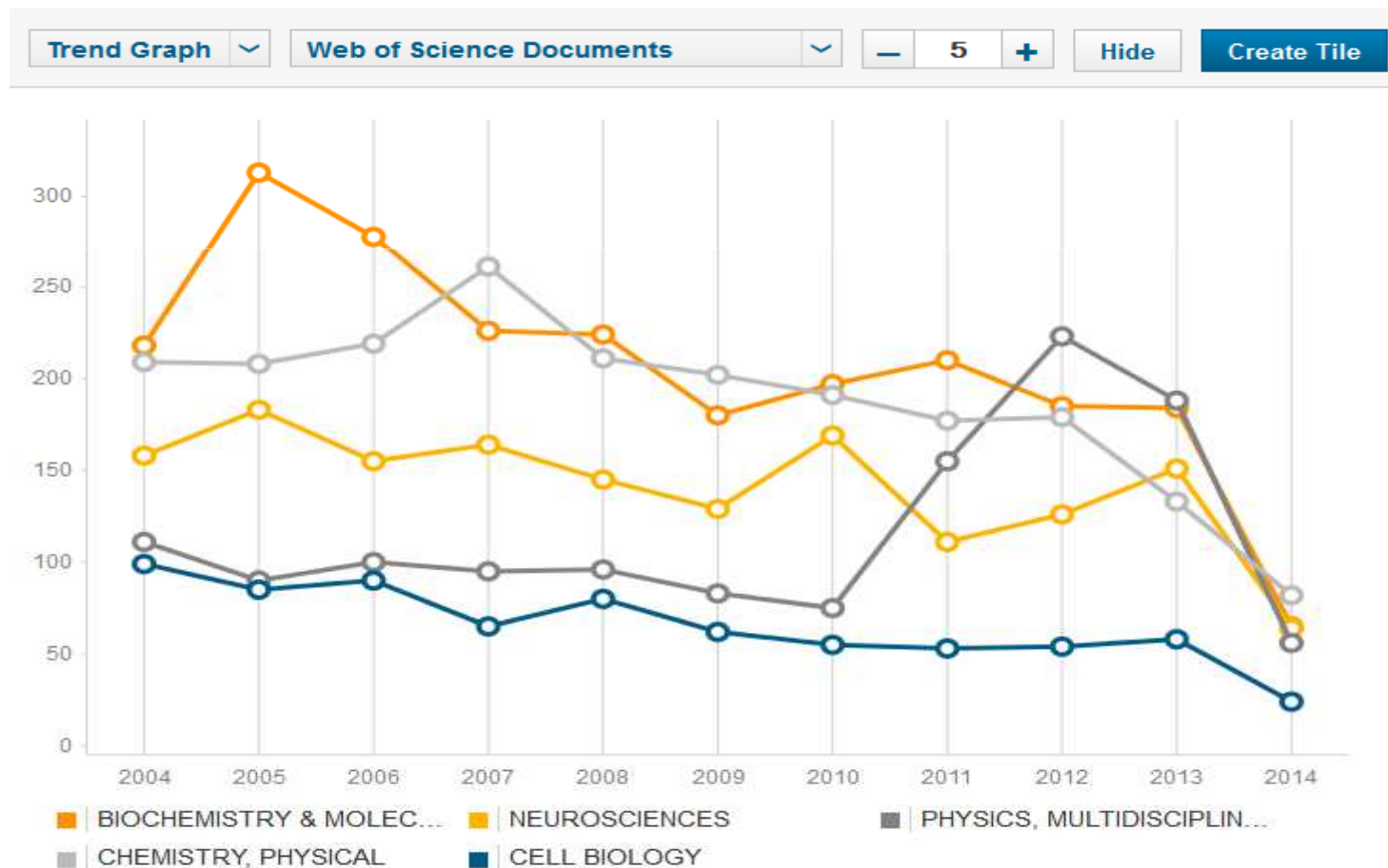
- Az MTA szakterületi teljesítménye idősorosan (kibocsátás), ESI





A mutatók érzékenysége a kiválasztott paraméterekre

- Az MTA szakterületi teljesítménye idősorosan (kibocsátás), WoS WCs





A mutatók érzékenysége a kiválasztott paraméterekre

- Továbbiak:
- Időablak-érzékenység:
 - Diakrón indikátorok: citációs ablak!
 - Historikus adatok
- Értelmezés:
 - Presztizs, teljesítmény, múlt, jelen stb.
- Nincs: Impakt faktor!!!
- Idézettségi Lelőhely-jellemzők (diakrón mérőszámok)



Az InCites használata: értékelés

- **Előnyei, használati értéke**
 - Javítja a kutatásértékelési kultúrát, szakmailag megalapozott mutatók széles körét kínálja
 - Szakszerű használata sokdimenziós értékelést tesz lehetővé
 - Az értékelés dimenziói átjárhatók, a szempontok kombinálhatók, tájékozódó, exploratív használatra ideális
 - Az mutatók az adatbázis mindenkori aktuális tartalmát közvetítik (frissítési gyakoriság függvényében)
 - Hatékony támogatást biztosít a tudománymetria művelőinek kutatásértékelési feladatok elvégzéséhez: szakszerű felhasználáshoz nagy segítség (benchmarkok)



Az InCites használata: értékelés

- **Veszélyei, kockázatai, gyengeségei**
 - Az adatminőség a közvetlen felhasználást kockázatosá teszi (hozzárendelés, aggregációs szintek)
 - A mutatók nagy választéka és értelmezése szakértelmet kíván
 - A lekérdezések rangsor-jellegű eredményt adnak, ami félrevezető (vö. a Scimago Institutional Ranking 2013-as médiavisszhangját)
 - Az eredmények komplett kutatásértékelési jelentésként kínálják magukat, noha bemenetet szolgáltatnak ehhez



Az InCites használata: értékelés

- **A jó gyakorlat elemei**
 - Tájékozódás, információ: kutatásmenedzsment
 - Rangsor-jellegű értelmezés elkerülése
 - Szakértői felhasználás: kutatásértékelés
 - Tisztított adathalmazok elemzése („Custom data?” vs. TTO)
 - Elemzést **támogató** szakmai háttérinformációk kinyerése (benchmarkok)





Jó gyakorlat

A KIK Tudománypolitikai és Tudományelemzési Osztálya
Konzultáció, szaktanácsadás és elemzések révén támogatja
az InCites szakszerű felhasználását!