

BIBLIOMETRIA

A *Library Trends* tematikus száma (Bibliometrics. Issue ed. POTTER, W. G., 30.vol. 1981. 172 p.) és LAUNI, S. M.: Bibliometrics. It's theoretical foundations, methods and applications c. írása (*Libri*. 31.vol. 1981. 4.no. 294–315.p.) BALÁZS SÁNDOR szemléje.

A *Library Trends* különszámában 15 szerző 11 tanulmánya, többszáz hivatkozással és gyakori ismétlődésekkel, szemleszerűen foglalkozik a bibliometria alapvető, főleg elméleti kérdéseivel. A *Libri* cikke túlnyomórészt alkalmazásra orientált, a gyakorlatból vett példákkal illusztrált helyzetkép.

Előzmények és dimenziók

A témakör első tanulmányának az *F. J. Cole* és *N. B. Eales* szerzőpártól 1917-ben megjelent összeállítást tekintik, amely az összehasonlító anatómia 1550 és 1860 között megjelent irodalmára terjedt ki.

1923-ban *E. W. Hulme* a statisztikai bibliográfia fogalmát használta a szóban forgó területre, de ez sok félreértésre adott okot, mert összetévesztették a statisztika bibliográfiájával.

1969-ben *A. Pritchard* vezette be a gyors elfogadásra talált bibliometria kifejezést,¹ ami meghatározása szerint:

„matematikai és statisztikai módszerek alkalmazása abból a célból, hogy megvilágítsa az írott kommunikáció folyamatait, valamint a tudományok természetét és fejlődését az írott kommunikáció különböző metszeteinek megszámlálása és elemzése útján”.

A bibliometriai módszerek népszerűségét és elterjedését igazolja a témakör legátfogóbb bibliográfiáiban közzétett dokumentumok száma:

A. Pritchard – 1969: 700 tétel,

Institute for Scientific Information – 1976: 500 tétel,

R. Hjerpe – 1980: 2032 tétel.²

Módszerek és modellek

A korai – és egyszerűségükben ma is gyakran alkalmazott – bibliometriai vizsgálatok *tapasztalati úton nyert* gyakoriság-eloszlásokat foglaltak táblázatokba és grafikonokba.

Többségükben egy változót hoznak összefüggésbe egy másik változóval. (Pl. folyóiratok vizsgálatánál két ilyen változó lehet: a lapok száma és a cikkek száma.) A nyert adatok táblázatba foglalásának két módja ismeretes.

A. Nagyság szerinti eloszlás alapján. Pl. S. C. Bradford (1934) egy vizsgálatánál:

n	$f(n)$	$nf(n)$
1	102	102
2	25	50
3	13	39
4	2	8
5	7	35
6	1	6
7	3	21
8	3	24
9	1	9
10	2	20
13	2	26
15	1	15
18	1	18
22	1	22
Összeg	$J = 164$	$N = 395$

Ennél: $f(n)$ azon folyóiratok száma, amelyek mindegyike n számú cikket közölt az adott szakterületen, ahol a vizsgált folyóiratok teljes száma $J = \sum f(n)$ és a megfigyelt cikkek teljes száma $N = \sum nf(n)$.

B. Rang szerinti eloszlás alapján. Az előző nagyság szerinti eloszlás alapján összeállított táblázat fordítottja, ahol prioritást kapnak a legtöbb cikket közölt („legtermékenyebb”) lapok.

r	$g(r)$
1	22
2	18
3	15
5	13
7	10
8	9
11	8
14	7
15	6
22	5
24	4
37	3
62	2
164	1

Itt $g(r)$ azoknak a cikkeknek a száma, amelyeket az r rangú (rangsorú) folyóirat közölt. A legtöbb cikket (22) közlő lap kapta a rangsor 1. helyét, a következő 18 cikkel a 2. helyet stb.

Ujabb bibliometriai vizsgálatokkal már megkísérik olyan modellek kidolgozását, amelyekben egy változó már két vagy több változó hatására válaszol. Ez a megközelítés meg fogja változtatni az eddigi általánosság³ definícióját, mivel az ilyen többváltozós modellek szükségszerűen magukba foglalják majd az eddig általános érvényűnek mondott korábbi egyváltozós modelleket is.

A bibliometriában nagy fellendülést hozott a számítógép alkalmazása, ami elsősorban a citátum-vizsgálatok és elemzések nagyarányú elterjedését tette lehetővé. Ezeknek alapja az a megállapítás, amit *J. M. Ziman* (1968) így fogalmazott meg: „*Egy tudományos publikáció soha nem áll önmagában, mindig bele van ágyazva a tárgykör irodalmába*”.

A citátum-vizsgálatok igen széles területén pontos különbséget kell tenni az alapvető fogalmak között:

hivatkozás: egy cikk hivatkozásainak száma azzal a számmal mérhető, ahány tétel szerepel irodalomjegyzékében, jegyzeteiben vagy rejtett bibliográfiájában, (angolul: *reference*);

hivatkozat: a hivatkozatok (hivatkozottság) száma attól függ, hogy egy cikk hányszor szerepel más cikkek irodalomjegyzékében, „citation index”-ekben stb., (angolul: *citation*);

citátum: a hivatkozások és hivatkozatok összefoglaló fogalma (az angol „*citation analysis*”-t sokszor ebben az összefoglaló értelemben használják).⁴

A gyakorlatban a citátumok elemzésénél – a legegyszerűbb összeszámláláson túl – két módszer használatos:

A bibliográfiai kapcsolódások vizsgálata. Két dokumentum bibliográfiailag akkor kapcsolódik össze, ha irodalomjegyzékükben egy vagy több azonos dokumentumra hivatkoznak.

Az együttes hivatkozatok vizsgálata. Két vagy több dokumentum akkor kapcsolódik ezen a módon, ha együttes hivatkozatként szerepelnek egy dokumentumban.

A bibliometriai módszerek között igen vitatott terület az *avulás*, ami *M. B. Line* és *A. Sandison* (1974) szerint „*az információ érvényességének és hasznosságának hanyatlása az időben*”. Ez a probléma széles körű érdeklődést nyert mind az elméleti, mind a gyakorlati szakemberek körében. Az avulás kritériumaiként általában a következőket vizsgálják: a dokumentumok használatának mértéke (olvasók, kölcsönzések stb.) és hivatkozataik „sűrűsége”. Sok kutató a „*lezési idő*” fogalmának alkalmazása mellett tör lándzsát.

Az irodalom avulását azonban sokan csak hipotézisként kezelik, mivel szerintük az információk érvényessége és a dokumentumok használhatósága között nincs közvetlen kapcsolat. Ezt igazolja, hogy sok avulási vizsgálat nem járt semmiféle eredménnyel. A *Library Trends* különszámában *D. K. Grapen* és *S. P. Milner* ezt a kérdést igen visszafogottan és kritikusan tárgyalva állapítja meg, hogy ezen a téren még széles körű alap kutatásra van szükség.

A bibliometria főbb törvényei

1. *A Bradford-féle ún. szóródási törvény.* *S. C. Bradford* törvénye annak a szabályszerűségnek a megfigyelésével kezdődött, ami a publikált irodalom keresésében mutatkozik meg. Általánosságban ez a különböző forrásokban rendszeres koncentrálttsággal vagy szétszórtsággal jelentkező információkat jelenti. Azaz pl. egy bizonyos téma keresésekor a releváns cikkek nagy száma kevés folyóiratban található, míg a továbbiak szétszórtan igen sok folyóiratban.

S. C. Bradford (1934) egy alkalmazott geofizikai bibliográfia tételeit vizsgálva a folyóiratcímeket három fő csoportba tudta osztani úgy, hogy mindegyikben nagyjából azonos számú cikket talált. Így:

az első csoport	9 folyóiratában	429 cikket
a második csoport	59 folyóiratában	499 cikket
az utolsó csoport	258 folyóiratában	404 cikket.

Ez az elrendezés egyúttal a folyóiratok csoportosításában megközelítette a következő szabályszerűséget:

9 folyóirat

$9 \times 5 = 45$ folyóirat

$9 \times 5 \times 5 = 225$ folyóirat

Nagyjából tehát 9-ben volt megállapítható a cikkek egyharmadát tartalmazó alapvető folyóiratok száma. A következő hasonló számú cikk megtalálásához már ötször annyi folyóirastra volt szükség. A harmadik csoportban pedig ismét ötször annyi folyóirat átvizsgálására volt szükség. Az alapvető folyóiratok számának és a szorzónak (amik esetenként különbözőek lehetnek) alapján a Bradford-törvény első elméleti megállapítása a következőképpen is kifejezhető:

$$1 : a : a^2,$$

ahol mindhárom csoportban azonos számú cikk található.

Ezt a törvényt később maga *Bradford* is továbbfejlesztette, de sokan mások is foglalkoztak kiegészítésével. *B. C. Brookes* (1968) olyan képletet dolgozott ki, ami nem függ a folyóiratok csoportosításától. Az eredményeket időlegesen befolyásoló tényezőket vizsgálva *W. Goffman* (1964) kidolgozta ún. „járványelméletét”. Más szakemberek ezeket az elméleteket egyéb területek matematikai modelljeivel vetették össze.

2. *A Lotka-törvény.* *A. J. Lotka* (1926) azt kívánta megállapítani, hogy a „különböző képességű emberek milyen mértékben járultak hozzá a tudományok fejlődéséhez”. Ebből a célból folytatott vizsgálatot, amely kiterjedt az 1907 és 1916 közötti „Decennial Index to Chemical Abstracts” 6891 szerzőjére és a „Geschichtstafeln der Physik”-ben 1900-ig felvett 1325 kiemelkedő fizikusra. A nagy és kis termékenységű szerzők eloszlását a tudományos irodalomban vizsgálva fogalmazta meg a produktivitás reciprok négyzetes törvényét. Megállapította, hogy „az n dolgozatot írt szerzők száma $1/n^2$ -e az egy dolgozatot írt szerzőkének... és az egy dolgozatot írt szerzők aránya az összes szerzőkhöz viszonyítva 60 százalék”. Azaz minden 100, egy adott időtartam alatt egy cikket közlő szerzőre 25 két-dolgozatos, 11 három-dolgozatos jut és így tovább.

Ugyanez más megközelítésben: a több mint n cikket írt szerzők száma az $1/n$ képlet szerint változik, ami azt jelenti, hogy 5 szerzőből 1-nek van 5 vagy több cikke; 10 szerzőből 1 legalább 10 közleményt írt stb.

Ezeket a megállapításokat 1949 óta idézik úgy mint a Lotka-törvényt.

Későbbi vizsgálatok több kritikai észrevétellel megkérdőjelezték ennek a törvénynek általános érvényességét. R. A. Fairthorne (1969) szerint Lotka alábecsülte a kimagaslóan termékeny szerzők számát, de törvénye jól alkalmazható a kevésbé termékenyekre. Többen csak megközelítésnek és nem szigorú eloszlásnak minősítették ezt az elvet. J. Vlachy (1972) széles körű vizsgálatokat folytatott az eltérő eredményeket kiváltó befolyásoló tényezők vonatkozásában.

Újabbán viszont többen elismerték, hogy amennyiben a vizsgálat időtartama minimum 10 év és a szerzők körét széles körben állapítják meg, akkor a Lotka-törvény igazolható. Az USA-ban pl. eredményesen használták az „Anglo-American Cataloging Rules” 2. kiadásának tervezésénél.

3. A Zipf-törvény. G. K. Zipf 1920 és 1949 között folytatott vizsgálatainak alapja, hogy minden nyelvben a szöveget igen kevés szó igen gyakori előfordulása jellemzi, míg a többi csak ritkán kerül felhasználásra.

Ennek a megkülönböztethető szó-előfordulásnak alapvető vizsgálata, amikor megszámloljuk az egy természetes nyelven írt szövegben előforduló szavak gyakoriságát. Ezután a szavakat előfordulásuk gyakoriságának megfelelő sorrendbe tesszük úgy, hogy a leggyakoribb szó kerül az 1. helyre, a következő a 2.-ba stb.

G. K. Zipf a sorrend és előfordulás közti arányt a következő képlettel közelítette meg:

$$rf = c$$

ahol r a szó sorrendje, f az előfordulás gyakorisága és c egy állandó (a szövegtesttől függően), általában a szövegben előforduló szavak számának egy tizede.

Grafikus ábrázolásban ugyanez a matematikai képlet:

$$\log r + \log f = \log c$$

R. E. Wyllis megjegyzése szerint a Zipf-törvény a bibliometria legtalányosabb jelensége. „A matematikusok azért hisznek benne, mert úgy tudják, hogy a nyelvészek saját területük törvényeként hozták létre – a nyelvészek viszont úgy fogadják el, mint a matematikusok törvényét.”

B. Mandelbrod (1954) az információlemélet alapján magyarázta a Zipf-törvényt, és mintegy első megközelítésként használta fel – a betűk és szóközök vizsgálata útján – a kommunikációs költségek csökkentésére.

4. Az exponenciális növekedés. A tudás exponenciális növekedésének felismerése – egyesek szerint – először Conan Doyle „The Great Keinplatz Experiment” című művében (1921) jelentkezett először. Itt írta, hogy „a tudás tudást szül, mint ahogy a pénz kamatot”. Eszerint a tudás növekedése a kamatos kamathoz hasonlítható, azaz matematikailag exponenciális.

D. J. D. Price (1963) tette széles körben ismertté a tudás, a tudományok és a tudományos irodalom exponenciális növekedését. Vizsgálatait a következő mutatókra alapozta: tudósok száma; tudományos lapok száma; tudományos publikációkról készült kivona-

tok száma; tudományra fordított összeg stb. A tudományos irodalomra vonatkoztatva megállapította, hogy annak növekedési üteme az elmúlt két évszázadban kb. 5 százalékos volt, azaz mennyisége 15 évenként megkétszereződött.

Később mind ő, mind pl. *D. Crane* (1972) és mások úgy találták, hogy ez az exponenciális növekedés lelassul(t), lineárisrá válik, ill. az exponenciális görbe logisztikussá alakul.⁵

Alkalmazások és példák

A bibliometriai vizsgálatokat céljuk és alkalmazásuk alapján csoportosíthatjuk. *Dokumentumok válogatása.* *J. C. Baughman* (1973) a „Social Sciences and Humanities Index” 24. évfolyamának – 1970–71 – „szociológia” és „társadalmi” címszavait átvizsgálva elemezte a cikkek hivatkozásait. A 11 130 hivatkozásból 6840 vonatkozott monográfiákra. Ezen belül 759 cím két vagy többszörös hivatkozással szerepelt, 98 pedig öttel vagy ötnél többel. Ezeket a műveket nevezte el úgy mint „néhány a szociológia legjobbaiból”.

S. M. Lawani a „Tropical Abstracts” négy évfolyamából (1967–70) a kivonatolt cikkek alapján rangsorolta a publikáló lapokat. A folyóiratok rangsorát állománygyarapítási útmutatóként használta.

E. Garfield (1975) a „Science Citation Index” ötmillió hivatkozását elemezve dolgozta ki a „hatásfok” fogalmát, melynek alapján egyes szakterületeken táblázatokban foglalta össze a nagyhatású folyóiratok jegyzékét.

Információkeresés. *G. Salton* (1963) szerint a citátumok vizsgálata jól kiegészíti a kulcsszavas/deszriptoros információkeresést.

I. Yermish (1975) interaktív számítógépes rendszere felhasználja a bibliográfiai rekordok citátumai közti kapcsolatokat. Ezt úgy éri el, hogy minden egyes dokumentum rekordját kiegészíti egy REFLIST (a dokumentum hivatkozásainak jegyzéke) és egy CITELIST (a dokumentum hivatkozatainak jegyzéke).

Több helyen bevált a bibliográfiai adatbázisok és a citátum indexek együttes alkalmazása a keresésben.

Gyűjtemények és információs szolgáltatások értékelése. Egyes könyvtárak állományuk értékelésénél jó eredményeket értek el a bibliometriai vizsgálatok során nyert rangsorok felhasználásával.

Egyes vizsgálati módszerek eredményeiből több helyen hasznos útbaigazításokat szereztek az állomány apasztásához.

Bibliográfiák, referáló és indexelő szolgáltatások értékelésére gyakran használt a *Bradford-törvény* és *Brookes-féle változata*. Ezek alapján felbecsülhető egy szakterület összes folyóiratainak száma, valamint a lapok összes cikkeinek száma. Így bizonyítható, hogy egy állítólagosan teljes bibliográfia vagy magát a szakterületen átfogónak nevező referáló/indexelő szolgáltatás valójában milyen mértékig az.

Szakirodalmi jellegzetességek. A jellegzetességek tanulmányozása és elemzése kiterjed a következőkre: kiadó országok; nyelv; hivatkozott kiadványok kora; dokumentum-

típusok használata; kapcsolatok más szakterületekkel; szóródás; a kommunikáció kritikus pontjainak meghatározása.

H. H. Fussler (1949) az elsők között disszertációjában a kémiai és fizikai irodalom használatát vizsgálta az USA-ban a következők szerint: más szakterületek irodalmának fontossága a kémia és fizika szempontjából; – a különböző dokumentumtípusok használata és relatív fontossága; – a legfontosabb sorozatcímek; – az USA-ban használt irodalom eredete országoként.

Avulás. A leggyakrabban vizsgált kérdés a következő paraméterek alapján: átlag életkor; – közép- vagy felezési idő; – évi öregedési tényező; – hasznossági tényező.

Ezek a vizsgálatok alapot nyújtanak a régi anyagok beszerzéséhez, a folyóiratok külső helyen való tárolásához, az állomány apasztásához stb.

Történelmi, szociológiai alkalmazások. A citátumok különféle kapcsolatainak vizsgálata igen hasznos a tudománytörténet tanulmányozásához. Ezek grafikai ábrázolásai (historiograph) rávilágítanak egyes tudományterületek kapcsolataira és elősegítik a fejlődés gyors áttekintését. Elősegítik a legfontosabb művek és események felismerését és azonosítását. Rávilágítanak arra, hogy bizonyos személyek, tudományos intézetek, események stb. miként befolyásolnak későbbi eredményeket.

H. Inhaber (1974) a tudományos szerzők és a lakosság számának összehasonlításával mutatta ki a különböző városok szakember-koncentrátságának mértékét.

A szociológiai alkalmazásnál kimutatható egyes személyek vagy kutatóhelyek stb. kiemelkedő szerepe.

A bibliometria oktatása

A. M. Schrader szerint a bibliometria oktatása igen elhanyagolt területe az információs szakemberképzésnek.⁶ Ezért – ajánlott irodalomjegyzékkel kiegészítve – erre a célra az alábbi (MLS diplomát nyújtó) posztgraduális tanfolyam tematikáját ajánlja:

- a) a terület áttekintése – 1 egység,
- b) elméleti alapok – 2 egység,
- c) kutatási tradíciók: törvények és modellek – 5 egység,
- d) kutatási tradíciók: empirikus leírások – 5 egység,
- e) alkalmazások a szakmai gyakorlatban – 2 egység, –

összesen 15 egység.

JEGYZETEK

1. Ujabban felmerült az „*informetria*” kifejezés is. Pl.: *Bonitz, M.*: *Informetria*. = Zentralblatt für Bibliothekswesen. 96.vol. 1982. 1.no. 18–22.p.
2. Kiegészítve a *Magyarországon megjelenő „Scientometrics”* 1982. májusi számában 518 tétellel 241–273.p.
3. Az egy változó hatását egy másik függvényében széles skálán vizsgáló általános modellek közül kiemelkedik *D. J. D. Price, A. B. Bookstein* és *B. C. Brookes* három modellje.

4. A magyar terminusok kérdésével csak 1977 után találkozhatunk. Pl.:

Ruff I. – Braun T.: A tudománymetria alkalmazása tudományágazati elemzésre I. = Magyar Tudomány. 1977. 2.sz. 117–125.p.

Száva-Kováts Endre: Az informatikai felezési idő. Bp. OMKDK, 1979. 403 p. (A tudományos tájékoztatás elmélete és gyakorlata 24.)

Braun Tibor – Bujdosó Ernő – Ruff Imre: A tudomány mint a mérés tárgya. Bp. MTA Könyvtára, 1981, 295 p. (Az MTA Könyvtárának informatikai és tudományelméleti sorozata 1.)

Száva-Kováts Endre: Az „Ortega-hipotézis” hivatkozatelemzéses cáfolata. = Tudományos és Műszaki Tájékoztatás. 1981. 8–9.sz. 337–358.p.

5. A bővebb ismertetés mellőzhető, mivel a magyar olvasó az exponenciális törvényre vonatkozóan átfogó képet kap *Price, D. J. D.:* *Kis tudomány – nagy tudomány.* Bp. Akadémiai K., 1979. 248 p. (Korunk Tudománya)6. Az *ELTE* információs szakán rendszeresen oktatnak bibliometriai ismereteket.

Hélène Desvals: Comment organiser sa documentation scientifique c. könyvének illusztrációiból

250 ÉVES fennállását ünnepelte 1981-ben a Philadelphiai Könyvtári Egyesület, az USA legrégebb kulturális intézménye. Alapítói – közöttük Benjámín Franklin – tagdíjából élő kölcsönkönyvtár létesítésére szövetkeztek. Könyvtáruk 1774–1800 között a nemzeti könyvtár funkcióját látta el. A 19. század végén, a modern közművelődési könyvtárak fellendülése idején úgy tértek ki a konkurenciaharc elől, hogy eladták a könyvtár 100 000 nem tudományos jellegű könyvét, árubból régi, ritka könyveket és kéziratokat szereztek; így alakult ki a ma is működő prézens tudományos könyvtár.

Zeitschrift für Bibliothekswesen und Bibliographie. 1982. 2.no.