

NICHOLAS, D. – RITCHIE, M.: Literature and bibliometrics. London, Clive Bingley, 1978. 183 p.

### Irodalom és bibliometria

A bibliometria ma is elfogadott definícióját PRITCHARD fogalmazta meg (1969). A bibliometria statisztikai vagy matematikai módszerekkel elemzi egy tudományterület, vagy egy tudományterület írott dokumentumainak természetét és belső törvényszerűségeit. A bibliometria a bibliográfiából nőtt ki, miután felismerték, hogy a szakirodalom növekedése és változásai tradicionális bibliográfiai módszerekkel nem értelmezhetők. A tradicionális bibliográfia és a bibliometria között az a lényeges különbség, hogy az utóbbi inkább mennyiségi mint diszkurzív módsze-

reket használ, az irodalom egészével s nem az egyes tételekkel foglalkozik.

A bibliometriai vizsgálatoknak két típusa van: az irodalom tulajdonságait *leíró* és az irodalom komponensei közötti viszonyt – az irodalom „*viselkedését*” – *feltáró* vizsgálatok. Teljes és pontos adatokat kell nyújtaniuk a gyorsan változó irodalompopulációról. A leíró jellegű kutatásokhoz tartozik a dokumentumok formai, tipológiai elemzése. A „viselkedés”-kutatások a különböző dokumentumtípusok, országok, nyelvek, témák irodalma közötti kapcsolatokat tárják fel. A témakapcsolatokra vonatkozik a legtöbb ilyen jellegű kutatás.

A szétszórtan végzett bibliometriai vizsgálatok, s néha nehezen érthető elméleti fejtegetések miatt egységes módszer és szemlélet még nem alakult ki. A könyv egy átfogó gyakorlati útmutató a bibliometriához, hangsúlyozva a könyvtári és informatikai alkalmazási területeket. Mellőzi a matematikai levezetéseket és a bonyolult elméleti érve-  
léseket.

A bibliometriát felhasználják a bibliográfiai kontroll javítására. Az irodalom mennyiségi növekedése hatással van a szekunder irodalom struktúrájára. A növekedési mérték kiszámítása és irányának előrejelzése jelentős segítség a kiadók számára ahhoz, hogy kialakítsák a szekunder szolgáltatások jövőbeli lefedési területét és összetételét. A megjelenési hely, téma, dokumentumtípus és nyelv szerinti megoszlási statisztikák hasznos információk a fordításigények előrejelzésére, az állomány vagy indexművek által le nem fedett fehér foltok felderítésére. A témák közötti kapcsolatok kimutatása dokumentumok osztályozásának analízisével, hivatkozás- vagy használatvizsgálatokkal kijelöli

egy szakterület szekunder szolgáltatásainak keretét. A hivatkozások idővizsgálata és az évenként megjelenő publikációmennyiség ismerete hasznos retrospektív bibliográfiák szerkesztéséhez. Az irodalomnövekedésre vonatkozó előrejelzéseket felhasználják raktári férőhely és szakemberszükséglet számításához.

A bibliometria információt szolgáltat a tudományos kutatás helyzetéről. Az egyes szakterületeken megjelenő irodalom mennyiségi növekedése vagy csökkenése jelzi a kutatás felfutását vagy visszafejlődését. Hivatkozás- és használatvizsgálatok feltárják a témák, országok és nyelvek közötti kapcsolatokat. Ezek alapján megállapítható, hogy a primer folyóiratirodalom egy adott szakterületen megfelel-e az olvasók igényeinek, vagy esetleg egy új – feltehetően interdiszciplináris – kiadványra van szükség. A bibliometria alkalmazható egy könyvtári részleg, vagy információs szolgáltatás megjavítására is.

A bibliometriai felmérésekhez szükséges adatok forrásai bibliográfiák, indexelő művek, könyvtári katalógusok, hivatkozások, nemzeti és nemzetközi statisztikai kiadványok. A bibliometriai tanulmány előkészítésekor mindig definiálnunk kell precíz bibliográfiai terminológiával a dokumentumoknak azt a csoportját, melyet vizsgálni akarunk.

Bibliometriai és bibliográfiai adatok között különbséget kell tenni az irodalom leírásakor. A fő különbség a funkció. A bibliográfiai adatok egy dokumentum azonosítását szolgálják, s olyan elemeket is tartalmaznak, amelyeknek nincs szükségszerűen bibliometriai értékük. Bibliográfiai adatnak csak akkor van bibliometriai értéke, ha valamit jelez az irodalom szerkezetéről. A bibliometria számára felhasználható adatok: szerző, cím, megjelenési adatok, címváltozások, kiadások gyakorisága, kiadás formája, megjelenés helye, nyelve, tárgykör, eredeti vagy kritikai mű, tartalom, méret, előfizetési ár, kölcsönzési forgalom, lefedettség a szekunder szolgáltatásokban, kapcsolódó információ. A felhasználható adatokból a vizsgálatokban csak azokat veszik figyelembe, amelyek könnyen hozzáférhetők. A jellemző adatok előfordulási gyakorisága és az együtt előforduló adatscsoportok típusai különbözőek tudományok, dokumentumtípus, kiadó, megjelenési hely szerint. Egyes tudományágak – pl. fizika, kémia – területén publikáló kutatók a gyors információközlési formákat, főleg a folyóiratot részesítik előnyben. A társadalomtudományok területén elterjedtebb a monográfia alapú kommunikációs rendszer.

## A bibliometriai vizsgálatokban felhasználható adatok értékelése

*Szerzőség.* A szerzőség három szempontból érdekes a bibliometriai kutatók számára. Egy szerző produktivitása jelzi tudományos tevékenységének intenzitását. Több-szerzős művek eloszlásvizsgálata tudományterületek szerint megmutatja, hogy melyik témákban általános a team-munka. A szerzők kiadók vagy folyóiratok szerinti megoszlás tükrözi az egy témával kapcsolatos publikációk tömörülését vagy szóródását kiadványok szerint.

A *kiadók* típusai három fő csoportra oszthatók: tudományos, kereskedelmi és intézményi. A kiadó és a megjelenő művek színvonala, típusa, megjelenési gyakorisága között szoros kapcsolat van. Különösen figyelemreméltó az a tény, hogy mindegyik tudományterületre jellemző egy kiadói típus, pl. földrajztudományban, pszichológiában

a kiadványok közel felét tudományos társaságok jelentetik meg. Országok szerint is eltérő, hogy melyik kiadó típus a legjelentősebb. Az USA-ban a felsőoktatási intézményeké, az NSZK-ban a kereskedelmi kiadóké a vezető szerep.

*Cím.* A cím az információcsere egysége. A megjelenő címek száma és növekedési aránya érdeklődik mindazokat a testületeket és egyéneket, akik szerszeményeznek, feldolgoznak és raktároznak.

*Megjelenés időpontja.* Egy mű többféle kiadásban is megjelenhet különböző időben, ez azt jelenti, hogy ugyanazt a művet tartalmazza preprint, reprint, monográfia, újabb kiadás. Egy műnek többszöri megjelenése növeli a kiadványszámot, ezek torzítják a kiadványtermés növekedési arányát.

A folyóiratok megjelenési és megszűnési időpontja nem jelzi egyértelműen egy kiadvány megszűnését vagy új indulását. Ennek oka a folyóiratok címváltozásai. A pontos folyóiratmennyiség megállapításához meg kell különböztetnünk a valódi megszűnést, amikor a folyóirat semmilyen formában nem folytatódik és a látszólagos megszűnést, ami címváltozás, több sorozatra bomlás vagy összeolvadás eredménye lehet. Ezeknek a változásoknak pontatlan meghatározásából adódik sok pontatlan bibliográfiai közlés.

Bibliometriailag a címváltozások a változások természete és a címek várható élettartama szempontjából érdekesek. A címváltozások lehetnek egyszerűek, amikor a címen kívül semmi sem változik. Ez a legáltalánosabb. Oka lehet:

- gazdasági, piacnövelő művelet (pl.: Library world – New library world)
- terminológiai változás (divatos, felkapott kifejezések gyorsan változnak)
- tartalmi bővülés vagy szűkülés (pl.: Library Science Abstracts – Library and Information Science Abstracts)
- kiadó változása
- kiadási forma változása (ami tulajdonképpen új típusú információtartalom közlésének következménye: pl. newsletter – journal)

**Összetett címváltozások típusai:**

- összeolvadás (a leggyakoribb ebben a csoportban; a beolvadónak ez a bibliográfiai megszűnése, a másik kiadvány változatlan, legfeljebb az alcím tükrözi a beolvadást)
- egyesülés (két vagy több kiadvány olvad egybe, egy teljesen új jön létre egyesítve az előzmények tulajdonságait)
- kiterjedés (egy folyóirat a témabővülés miatt több részre szakad).

A címváltozás gyakori jelenség, időtartama szakterületenként eltérő. Építészeti folyóiratoknál 7–8 év, pedagógiai folyóiratoknál 14 év. Az adatgyűjtés a címváltozásokról teljesen megoldatlan. Egyes könyvtárak katalógusai közlik a címek teljes életét.

*Megjelenés gyakorisága (folyóiratok).* A megjelenés gyakoriságát a termelt információmennyiség, a kurrens információ iránti igény, a kiadás gazdaságossága szabályozza. A kiadványszám nem azonos a kiadás gyakoriságával. A kiadók két számot egyszerre is megjelentethetnek. Ekkor a gyakoriság kisebb, de a kiadványszám ugyanaz.

*Publikálási formák.* Különböző típusú dokumentumok különböző információfajtákat hordoznak. Egyes formák csak egy típusú információt tartalmaznak ilyenek az abstract-ek. A folyóiratok, évkönyvek, újságok, többféle információt tartalmaznak.

A kiadványtípusok előfordulása szakterületenként változik. A bibliometriai analízisek a szakirodalom természetét a leggyakoribb dokumentumtípus, a folyóirat alapján végzik. Más típusok – könyv, szekunder kiadványok – természetének, kapcsolatának kutatása még nem kezdődött meg.

*Kiadás helye.* Bibliometriai szempontból a megjelenés helye a kiadvány nemzetiiségét jelenti. Egy szakterület nemzeti irodalomtermése tükrözi a nemzeti kutatás, kiadói terület fejlettségét, szerzők alkotóképességét. A bibliometria vizsgálja, hogy milyen nyelveken jelenik meg egy-egy tudományterület szakirodalma és melyik nyelvnek milyen jelentősége van az egyes tudományterületeken; milyenek a fordítási szokások és milyen a nyelvek közötti kapcsolat. A kapott adatok alapján szakterületenként meghatározható, hogy milyen nyelvű publikálással biztosítható az információ legszélesebb körű elterjesztése. Fordítóirodák, kutatók nyelvoktatásának megszervezéséhez is fontos segédanyag.

A monográfiák zöme négy nyelven íródik: angolul, franciául, németül és spanyolul. A folyóiratirodalom zöme angol, orosz, német, francia és japán nyelven jelenik meg. A megjelenő művek nyelve és a megjelenés helye között nincs mindig összefüggés. A kis tudományos felvevő piaccal rendelkező országokban gyakran publikálnak valamelyik világnyelven.

*Téma.* A bibliometriai analízisek témaorientáltak. Bibliometriai szempontból kétféle szakirodalomról beszélhetünk. Az első típust az elfogadott tudományrendszerek szerint, a másikat a használat alapján definiálják. Az első egy konkrétan körülhatárolt témakör, a másik határai kevésbé körülírtak, s változnak az egyéni felfogás szerint. Pl. az első alapján vizsgáljuk a szociológia irodalmát, a másik alapján azt az irodalmat, amit a szociológusok használnak.

Mivel a tudományterület meghatározása eltérő lehet a felmérést végző nemzetisége, iskolai végzettsége, egyérfi orientációjától függően, a vizsgálat eredményeinek értékelésekor figyelembe kell venni ezt a bizonytalansági tényezőt.

*Eredeti vagy kritikai mű.* Az irodalom lehet a kutatás nyersanyaga (eredeti szövegek) vagy kritikai (róla szóló) színvonalra szerint gyermek, népszerű, tudományos irodalom. Adatokat ezekre a jellemzőkre egyes bibliográfiákban lehet találni.

*Tartalom.* A dokumentumok tartalmával a bibliográfiai leírások nem foglalkoznak. Monografikus kiadványoknál a cím többnyire eligazít a tartalomról. Más dokumentumok – legjellemzőbb típus a folyóirat – többféle információt tartalmaznak. A legtöbb folyóiratban a cikkeken kívül szekunder szolgáltatásokat is közölnek. Ha a szekunder irodalom elemzésekor a forrásokat mellőzzük, nem kapunk a valóságnak megfelelő képet. Ha az információforgalom természetét és mennyiségét akarjuk vizsgálni, a dokumentumok tartalmával is kell foglalkozni. A gyakorlatban a következő információ-„csomagokat” lehet megkülönböztetni: folyóiratcikk, hírek, ismertetés (review), hivatkozások, referátumok, konferencia-anyagok, címjegyzékek, kalendáriumok, levelek, hirdetések, bibliográfiai szolgáltatások. Bibliográfiai forrásokban a dokumentumok tartalmára vonatkozó adatok ritkán fordulnak elő, esetleg csak annyit közölnek, hogy a mű tartalmaz bibliográfiát.

*Árak.* Az árak változnak dokumentumtípus és szakterületek szerint. Az utóbbi évekre jellemző az állandó áremelkedés. Ár-adatok statisztikai kiadványokból nyerhetők.

Folyóiratoknál gyakori, hogy többféle árú van az előfizető könyvtár státusától függően. Ár-adatokat a bibliometriai költség/haszon vizsgálatokban használják fel. A felmérések polcméter/ár, hivatkozás/ár számításokat tartalmaznak. Ezek a felmérések kimutatják, hogy az olcsó, de ritkán használt folyóiratok többre kerülnek, mint a népszerű de magas előfizetési árú folyóiratok.

*Kölcsönzési adatok.* Ezeknek az adatoknak az alapján egy adott szakterület olvasótáborának mértékét becsülik.

*Lefedés szekunder szolgáltatásokban.* Ha egy folyóirat több szekunder szolgáltatásban előfordul, jelzi az adott cím fontosságát, interdiszciplináris természetét.

*Az irodalom mennyisége és növekedési aránya.* Különbséget kell tennünk az információ-mennyiség és az ismerethordozók számának növekedése között. A mű lapszáma nem jelenti az ismeretanyag, az új információ mennyiségét. Ezek a bibliográfiai adatok csak részlegesen nyújtanak helyes információt a felhasználás céljától függően. Állománygyarapítás és katalogizálás szempontjából a címek száma elegendő, de raktártervezés szempontjából már a dokumentumok vastagságát is jó lenne ismerni.

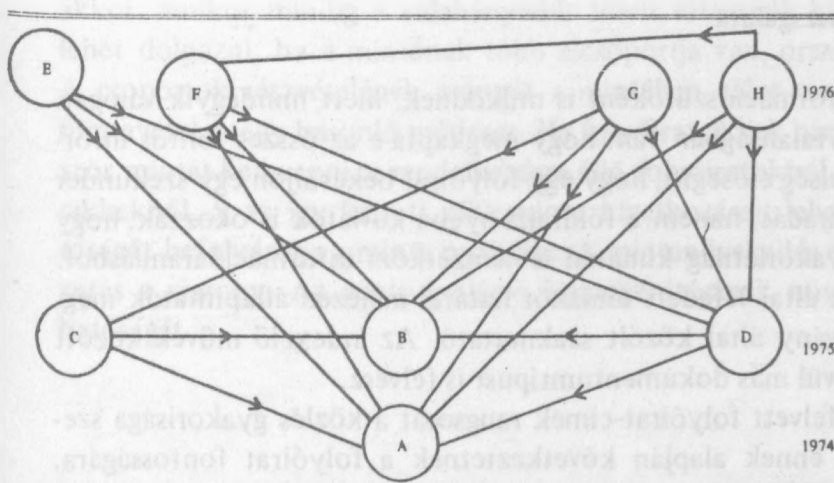
Az irodalomnövekedés mértékének problémájára már a század elején több kutató felfigyelt. A hatvanas években PRICE mérései alapján az egész könyvtáros-társadalmat foglalkoztató problémává vált. Szerinte az irodalommennyiség exponenciálisan növekszik és 10–15 év alatt megkétszereződik. A jelenlegi vizsgálatok kimutatták, hogy az exponenciális növekedés csak egy hirtelen fejlődési időszakra jellemző. A statisztikai modellek alkalmazásának helyességét az irodalomnövekedés becslésére többek között ROSE (1972) is kétségbe vonta. SANDISON (1974) új mérési eljárásokat javasolt. Az irodalomnövekedés mérhető a megjelent címek száma alapján. Ebben az esetben azonban a több művet tartalmazó címek egy tételként jelentkeznek. A folyóiratcikkek számlálásával is csak hiányos adatokat kapunk, mert nem minden tudományterületen ez a közlésforma a legelterjedtebb. Az oldalszám több szempontból előnyösen alkalmazható az irodalommennyiség növekedésének mérésére. A lap kevésbé változékony egység mint a cím, vagy a cikk, a lap és szavak száma között összefüggés van, könnyen hozzáférhető adat. A dokumentum vastagsága is egy lehetséges mértékegység. A vastagság (vagyis az elfoglalt polchossz) arányos a cikkek számával és az oldalszámmal. Ennek ellenére nem alkalmazható általános mértékegységként, mivel a kiadványok különböző vastagságú papírból, különböző kötéstípusokkal készülnek.

## Kapcsolatok, társulások, csoportok

A dokumentumok fizikailag önálló egységek, de szellemileg egymással kapcsolatban vannak. Minden szerző valamilyen módon foglalkozott mások munkáival és azok eredményeit beépítette a sajátjába. *Az irodalom ezért nem dokumentumok véletlenül egymás mellé került csoportjaiból áll, hanem kapcsolatok, csoportok állandóan mozgásban levő hálózatából.* Három fő kapcsolat kialakítási mód van. A hivatkozások kapcsolatot létesítenek az adott mű és több más mű szerzője között; indexművek a primer és szekunder

dokumentumok között; egy dokumentumon belül egymásmellettség is kialakít kapcsolatot. Hivatkozásanalízisek, könyvtárhasználati vizsgálatok teszik lehetővé a kapcsolatok felderítését.

*Tárgykörök közötti kapcsolat.* A témák közötti kapcsolatokat hivatkozásvizsgálattal tárják fel. Hivatkozásvizsgálattal megállapítható, hogy egy téma kutatása során milyen mértékben használják fel más ismeretágak irodalmát, vagyis milyenek az interdiszciplináris kapcsolatok. A tudományágak közötti kapcsolat általában kétoldalú, tehát kölcsönösen hivatkoznak egymásra. Vannak azonban egyoldalú hivatkozások is. Kimutatták, hogy a kriminológiai irodalomban nagy a hivatkozások száma a szociológiára, fordítva azonban ez nem érvényes. Azok a tudományágak információ-exportálók, amelyekre többször hivatkoznak, mint ahogy azok hivatkoznak másokra. Azok a tudományágak információ-importálók, amelyek többet hivatkoznak, mint amennyire rájuk hivatkoznak.



Az egymással kapcsolatban álló dokumentumok közötti viszonyt grafikusán ábrázolhatjuk. A dokumentumok összefüggő hálózata a klaszter. A kapcsolatok erősségét és irányát is lehet ábrázolni a klaszterben (1. ábra).

Amikor két cikk közösen hivatkozik egy vagy több műre, akkor a cikkek között bibliográfiai kapcsolat van. Mennél több cikkre hivatkoznak közösen, annál erősebb a bibliográfiai

kapcsolat, annál valószínűbb, hogy a két cikk hasonló témáról szól.

Azt az esetet, mikor két korábbi dokumentumra hivatkozik egy későbbi dokumentum, együtt-hivatkozásnak (co-citation) nevezzük. Az együtt-hivatkozás megmutatja, hogy a szerzők egy szakterületen mely folyóiratot tekintenek fontosnak. (Ld. az 1. ábra A és B dokumentumra vonatkozó hivatkozásait.) Az együtt-hivatkozás – mivel függ a hivatkozó szerzőktől – időről időre változik. A hivatkozás-kapcsolatok behatárolják a tárgykört, meghatározzák a kulcsszerzőket és -folyóiratokat.

*Ország és nyelv kapcsolatok.* A legtöbb tudományág kutatásában általában néhány országnak van vezető szerepe. A hivatkozások földrajzi szóródását az egyes országokban termelődő információmennyiség határozza meg. Az USA irodalma áll első helyen a más országokban megjelent dokumentumok hivatkozásmennyiségét tekintve.

*Avulás.* Sok vita tárgya az az időszak, amely folyamán egy dokumentumra hivatkoznak, vagy a könyvtárban még használják. A használatcsökkenés mértékének ismerete meghatározza a selejtezési, köttetési, hiánypótló szerzeményezési politikát, bibliográfiák

által lefedett időszak megválasztását. LINE (1974) felülvizsgálta az avulással kapcsolatos feltételezéseket és megállapította, hogy a gyors avulásnak különböző okai lehetnek: az adatok efemer jellegűek, új formában ismét megjelennek, gyorsan fejlődik az adott ismeretág. Az avulás lassabb a monografikus, leíró művek, koncepciókat ismertető elméleti és kritikai művek esetében. A legfrissebb irodalom kiemelkedő mértékű használatát PRICE (1963) szerint a kétféle irodalomhasználat – az újjal lépést tartó témafigyelés és a konkrét témára vonatkozó irodalomkeresés – egybeesése okozza. SANDISON (1974) a fizikai folyóiratok használatának értékét korrigálta az irodalomnövekedési adatokkal, s így szerinte a két évnél régebbi folyóiratok használatában nem mutatkozik lényeges csökkenés. A dokumentumtípus is meghatározó a használat változásában. A bibliometriai vizsgálatok főleg a folyóirathasználat elemzések alapján általánosítottak. A megállapítások nem érvényesek azokra a tudományterületekre, ahol nem a folyóirat a legfőbb kommunikációs csatorna.

### **Indexművek hatékonyságának vizsgálata**

Ezek a szolgáltatások információszűrőként is működnek, mert mindegyik válogat. Emiatt a használó örök bizonytalanságban van, hogy megkapta-e az összes fontos információt. A földrajzi, nyelvi közelség elősegíti, hogy egy folyóirat bekerüljön egy szekunder szolgáltatásba. Nemcsak a lemaradás, hanem a földrajzi-nyelvi korlátok is okozzák, hogy a harmadik világ irodalma gyakorlatilag kimarad a nemzetközi információáramlásból. Egy-egy szekunder szolgáltatás által lefedett témakör határai nehezen állapíthatók meg. Az egyetlen útmutatás a kiadvány által közölt szakmutató. Az indexelő művek között ritka, amelyik folyóiratokon kívül más dokumentumtípust is felvesz.

Az indexelő művekben felvett folyóirat-címek rangsorát a közlés gyakorisága szerint meg szokták állapítani s ennek alapján következtetnek a folyóirat fontosságára. A rangsorolás során nem tesznek különbséget a folyóiratok különböző tartalmi felépítése és periodicitása között. Ezért a gyakran megjelenő, több rövid cikket közlő folyóiratok mindig a rangsor elejére kerülnek. Az indexművek időszerűsége nem éri el a műfaj által megkövetelt szintet.

A különböző szolgáltatások közötti átfedések – mint a pazarlás egyik formája – az utóbbi idők gazdasági problémái miatt csökkennek. GILCHRIST és PRESANIS (1971) elemezte a Library and Information Science Abstracts, Referativnűj zszurnal és Information Science Abstract átfedését, és 706 címből 51 volt közös mindháromban.

### **Egy bibliometriai tanulmány megtervezése**

Elsőnek az adatgyűjtést kell megszervezni. Az adatok listáját egy speciális rendbe szervezik. Ez lehet alfabetikus, időrendi stb. Az adatgyűjtési időt a lehető legrövidebbre kell redukálni, mert így csökken a hibaszázalék. Az összegyűjtött adatok értékelése sokféleképpen végezhető, de van egy pár általánosan előforduló feladat. Az adatnagyság

megszámlálása is értékes eredményekhez vezet. Az adatok közötti összefüggések felderítése után kerül sor a kutatási jelentés megírására. Az adatok mellékletként való közlésével lehetővé válik, hogy az olvasók esetleg más módon is interpretálják azokat. Az adatok ábrázolására a legáltalánosabb eszköz a táblázat, de hatékonyabb az adatok grafikus vagy képi ábrázolása. Különböző diagramok tartoznak ide, kördiagram, oszlopdiaagram, hisztogram stb.

## Néhány megjegyzés a mintavételről

Az egész populáció minden egyes tétele nem vizsgálható meg. Az irodalomban érvényesülő törvényszerűségek lehetővé teszik, hogy megfelelően kiválasztott minták alapján vizsgálják az egészet. A mintát többnyire szekunder szolgáltatásokból lehet nyerni. Többféle mintavételi módszer van. Véletlen minta az, amikor a mintavétel populáció bármely tagja kiválasztható. Félig véletlen, vagy szisztematikus mintavételről beszélünk akkor, amikor mindig a valahányadik tételt választják ki. Rétegzett mintavétellel akkor lehet dolgozni, ha a mintának több alcsoportja van, ország, tárgykör, nyelv stb. szerint. A csoportok részvételének arányát a mintában előre meg kell határozni. A többlépcsős mintavétel is egy hasonló módszer. Ha folyóiratcikkek hivatkozásait vizsgáljuk, akkor először mintát kell venni a rendelkezésre álló folyóiratokból, majd a folyóirat-mintában levő cikkekből. S az így kapott cikk-minta hivatkozásait lehet elemezni. A minta megbízhatóságát befolyásolja a minta nagysága. A mintanövekedés és a megbízhatóság között négyzetes a viszony. Az adatpopuláció homogenitásának növelése is javítja a minta megbízhatóságát.

TÖRÖK Sára