

## **VI. A DOKUMENTUM- ÉS TARTALOMELEMZÉS MÓDSZERTANA** (Dr. Kiss Virág Ágnes)

### **Bevezetés**

Számos esetben a szekunder források nagyszámban érhetőek el a kutatási témában, azonban a dokumentumelemzés mégsem tárgyalt a hazai szakirodalomban olyan részletességgel, mint a kérdőíves megkérdezés, a mélyinterjú vagy a többi, jelen könyvben is szereplő módszertan. A dokumentumelemzés kiváló alapot adhat a kutatások elkészítéséhez, maga a módszertan az igényes szekunder forrás kutatás alapja is, ezáltal segít megalapozni a további primer kutatásokat, azonban a megfelelő mélységű és kidolgozottságú dokumentumelemzés önálló primer kutatásként is megállja a helyét.

### **1. Mi a dokumentumelemzés?**

A dokumentumelemzés egy olyan kutatási eszköz, amely során írásos tartalmakat elemzünk, értelmezünk annak érdekében, hogy valamilyen jelenség vagy esemény hátterét jobban megértsük, üzeneteket vagy rejtett információkat találjunk retrospektív módon. Többféle dokumentum összevetése és elemzése új összefüggéseket, mintázatokat vagy ötleteket tárhat fel akár egy vállalkozás működése vagy egy gazdasági jelenség kapcsán, illetve összetett, bonyolult intézményrendszerek működéséről, folyamatairól jól elemezhető információk gyűjthetők össze és rendszerezhetők (Jakusné, 2023). Bowen (2009) szerint a dokumentumelemzés dokumentumok szisztematikus áttekintését és értékelését foglalja magában, így alkalmas lehet mind a szakirodalmi áttekintés elkészítésére a kutatás kezdeti szakaszában, továbbá a probléma meghatározásában vagy a hipotézisek kialakításában, de primer kutatásra is, például egy szervezet, intézmény működésének elemzésére, egyes események kontextusba helyezésére, esettanulmányok készítésére vagy a feltárt adatok validálására, ellenőrzésére (Jakusné, 2023).

A dokumentumelemzés számos tudományágban elterjedt és alkalmazott, például szociológiában, gazdaságtörténetben, jogtudományban vagy marketingkutatásban. A dokumentumelemzés tehát elsősorban a szövegben megjelenő, a kutatás szempontjából releváns tartalmakat emeli ki. Ezzel szemben a tartalomelemzés már a szöveg mögé tekint, elemelve a szavakat, szókapcsolatokat, a jelzők használatát, a logika felépítését, és segít a kontextus alaposabb megértésében, árnyaltabb háttér kialakításában, feltárja a rejtett tartalmakat (Cserné – Ponyi, 2018). Tehát jellemzően a dokumentumelemzés részét képezi a tartalomelemzés.

### **2. A dokumentum- és tartalomelemzés forrásai**

A dokumentumelemzés tárgya lehet bármilyen írásos, képi, művészeti vagy audio anyag, küldetés- vagy sajtónyilatkozat, statisztikai riport, vezetői összefoglaló, reklám, rádiófelvétel stb. Például egy vállalat kapcsán vizsgálhatóak a megállapodások; üzleti tervek; üzleti folyamat- vagy interfész-dokumentációk; üzleti szabályok gyűjteményei; az aktuális folyamatok leírásai; marketingiratok; problémánapló (issue log); eljárások, irányelvek, előírások; egyéb jogi szabályozói dokumentumok, mint például törvények, előírások, rendeletek stb.; ajánlattételi felhívás (request for proposal); és használati esetek (use cases), jelentések, projektbeszámolók, szervezeti dokumentáció – szervezeti és működési szabályzat,

intézményi minőségirányítás program (PMI, 2020; Cserné – Ponyi, 2018). A dokumentumelemzés tárgyaként megjelenő anyagok nem feltétlen állnak szoros kapcsolatban a kutatás pontos tárgyával, de a kutató teremthet indokolt esetben kapcsolatot ezek között (Cserné – Ponyi, 2018). Ez utóbbira példa lehet egy rádióinterjú vagy egy film-, reklám- vagy könyvrészlet, ami alátámasztja például a korszellemet vagy a társadalmi kontextust.

Csoportosításukat tekintve lehetnek hivatalos dokumentumok, melyek a nyilvánosság számára is elérhetőek, annak címzettek (statisztikák, sajtóanyagok, honlaptartalom); személyes (naplók, levelek, feljegyzések) (Cserné – Ponyi, 2018) vagy belső dokumentációk (vállalati fehér könyv, működésre, jógyakorlatra vonatkozó anyagok). További dokumentumok lehetnek az úgynevezett személyiséget előhívó tartalmak, amelyek a vizsgált egyén vagy szervezet magánszféráját, belső szervezeti vagy kulturális működését tárják fel, például egy interjúban vagy dolgozói beszámolóban. Ezenkívül vannak a hétköznapi megnyilvánulásokról fakadó egyéb, plusz információk, mint a szókinccs vagy stílus, amit az egyén vagy szervezet képvisel (Cserné – Ponyi, 2018).

### ***3. A dokumentum- és tartalomelemzés általános módszertana***

A dokumentumelemzés egyik legfontosabb része ahhoz, hogy tudományos szinten is elfogadott, megbízható eredményt mutathassunk be, a pontos prezentálása annak, hogy milyen logikai menet, milyen kiválasztási szempontok alapján kerültek a dokumentumok, tartalmak a kutatásba. (Ezt a szekunder források kapcsán a szakirodalmi áttekintést megelőzően vagy az értekezés Anyag és módszertan fejezetében is érdemes bemutatni, szemléltetve, mely szempontok, modellek vagy logikai menet alapján kerültek az egyes irodalmi források vagy témakörök kiválasztásra.) Ismertetni kell, hogy az adott dokumentáció mikor készült, ki készítette, honnan származik – például nyilvánosan elérhető adatbázisból, céginformációs jelentésből, vállalati belső dokumentációból vagy magángyűjteményből stb. A dokumentumokat összegyűjtésük után strukturálni kell, hogy a megfelelő rendszerben lássuk a kapcsolódó tartalmakat, majd érdemes tartalomelemzést végezni rajtuk.

Krippendorff (1995) alapján aktualizálva a tartalomelemzés a következő folyamatban írható le (Cserné – Ponyi, 2018 alapján átszerkesztve):

*Adatkészítés* – azaz a szövegek vagy azok egyes egységei adatokká válnak, digitalizálódnak. Ehhez szükséges lépés az *egységmeghatározás*, azaz a kóddá vagy adattá váltás meghatározása, amik lehetnek szavak, mondatok, bekezdések vagy más egységek, gyakoriságok. Például visszatérő szavak – hányszor jelenik meg a fenntarthatóság kifejezés, vagy hány cikk jelent meg a témában, hányszor következett be egy esemény – piaci vezető helyzet vagy likviditási gond, esetleg rendeztek meg egy eseményt. Ez a kutatási témától függ.

*Mintavétel* – aminek keretében kiválasztjuk a témához, a kutatási kérdés megválaszolásához szükséges (kiinduló) forrásokat. Előfordulhat egy-egy forrásban más dokumentumra történő utalás okán érdemes újabb tartalmakat is beemelni a kutatásba. Pontosan dokumentálni kell a kiválasztott források eredetét, bekerülési okát.

*Kategóriák vagy kódok meghatározása* – pontosan meghatározzuk a szabályokat, melyek alapján az egyes egységek besorolásra kerülnek. Például kategória: külső jellemzők – minden olyan említés idekerül, melyben egy adott fejlesztés külső design szempontjából jelenik meg; funkció – a fejlesztés minden használati, alkalmazási tulajdonsága, ami a rendeltetés szerinti működéshez kapcsolódik stb. Ezután szükséges az:

*Adatredukció* – azaz a nem lényeges, irreleváns tartalmak kiszűrése, elhagyása, hiszen nem minden részlet szükséges a kutatási téma kapcsán, ami a dokumentumokban megjelent. Ezenkívül a sok elágazás nehezíti az alapos feltárást, konkrét irányok és következtetések megfogalmazását.

*Következtetés* – meg kell határozni, hogy a kinyert adatok alapján mi mondható el a vizsgált jelenségről, alanyról, azaz a kutatásunk tárgyáról.

*Elemzés* – analizáljuk, hogy a kutatott jelenség és a tartalomelemzés során nyert eredmények között milyen összefüggés van.

*Eredmények bemutatása* – a tartalomelemzés disszeminációjához bármely kvalitatív módszer eredményeinek szemléltetése alkalmas lehet (gondolatterkép, diagram, szófelhő), azonban a számszerűsített vagy több szöveges tartalmat bemutató szemléltetéshez érdemes táblázatokat is használni.

#### **4. Bibliográfiai elemzés**

A bibliográfiai (vagy bibliometriai) elemzés a tudományos szakirodalom kvantitatív alapú elemzését takarja, mely alapján az egyes szakterületek mintázatait, trendjeit és hatásaikat tárja fel statisztikai adatokon keresztül (Passas, 2024). Kvantitatív megközelítéséből fakadóan jóval objektívebb adatok képest adni a tudományos tudás alakulásáról, mivel megvalósul benne a szisztematikusság, az átláthatóság és a reprodukálhatóság, csökkentve a torzítás esélyét. Kiváló eszköz lehet a szakirodalom és a saját témánkhoz kapcsolódó kutatások előzetes vizsgálatához, a lényegesebb források, kulcsszavak felismeréséhez és az új irányzatok azonosításához is.

A bibliográfiai elemzések népszerűsége ma már elvitathatatlan, hála a nagy adathalmazok kezelésére is alkalmas szoftvereknek (R csomag, VOSviewer, Biblioshiny vagy az ingyenes, online Vöyant Tools), illetve a tudományos adatbázisoknak, mint a Google Scholar, Scopus, Web of Science stb.) (Passas, 2024). Felsorolásszerűen bemutatva elmondható, hogy a módszer alkalmas az alábbiakra (Passas, 2024):

- teljesítménymérés (performance analysis) – legtöbbet publikáló, legtöbbet idézett szerző, cikk, könyv egy terület vagy kérdés kapcsán,
- tudományos szerkezet vizsgálatra (science mapping) – együttműködési hálózatok (szerzőtársak szerkezete), hivatkozási hálózatok, tematikus kapcsolatok (milyen kulcsszavak vagy témakörök fordulnak leggyakrabban együtt elő).

Ezen módszerek segítségével feltérképezhetőek egy-egy tudományág vagy kutatási terület főbb témakörei, trendjei, együttműködési mintái, kiemelkedőbb figuráira.

Az adatgyűjtés megtervezése és végrehajtása az alábbi lépések mentén javasolt (Passas, 2024):

1. *Célkitűzések és központi kérdés meghatározása*: pontosan tudni kell a gyűjtés előtt, mely problémára keresünk választ vagy milyen jellegű elemzést szeretnénk csinálni (például struktúrakutatás, trendazonosítás, együttműködés vizsgálat). Ezen konkretizálás keretét ad a kutatásnak, segít a fókusz megtartani a munka során. Ennél a lépésnél még jellemzően nincs szükség szoftveres támogatásra, bár egyre jellemzőbb a mesterséges intelligencia előzetes megkérdezése a megfelelő is a kutatási kérdés(ek) pontosításához, de a megfelelő előzetes tudás, a témában való jártasság nagyon fontos lehet.

2. *Adatforrások kiválasztása:* ki kell választani, hogy milyen adatbázisok mentén gyűjtjük az adatokat. Lehetőség szerint minél nagyobb adatbázisokat alkalmazzunk, esetleg több, különböző forrású tudományos adatbázist együtt a releváns irodalom lekérdezéséhez (WoS, Google Scholar, Scopus, Science Direct) (Passas, 2024). Itt fontos kiemelni, hogy természetesen ez akkor javasolt, hogy a teljes tudományágat vizsgáljuk, illetve ha természettudományokban, élettudományban készült kutatásokat keresünk, mert azok nagy többsége a nemzetközi platformokon érhető el, de a bölcsész- és társadalomtudományokban szükséges további hazai adatbázisok beemelése (Cabezas-Clavijo – Torres-Salinas, 2021). Továbbá, ha csak a hazai irodalmat vizsgálánk, akkor szűkíthetünk csak azon adatbázisokra.
3. *Keresési stratégia és lekérdezés:* Lényege, hogy a téma ismerete mentén határozzuk meg a kulcsszavakat, szinonimákat és logikai operátorokat (AND, OR, NOT), amikkel a téma leginkább lefedésre kerül. Meg kell határozni a lekérdezési intervallumot (megtől meddig készült publikációk kellenek – nagyobb, hosszú idősoros elemzés a cél vagy az aktuális fejlemények) vagy az egyéb szűrőket (nyelv, dokumentumtípus – könyv, folyóirat, konferenciaközlemény stb.). A legtöbb esetben már van egy előzetes irodalomkutatásunk, melyből a legfontosabb keresőkifejezések nyilvánvalóvá válnak. Amennyiben több adatbázisból dolgozunk, külön-külön futtassuk le a gyűjtést, majd egyesítsük az eredményeket. A találatokat bibliográfiai adatformátumban (RIS, BibTex, CSV – ezeket fel is ajánlják a nagy oldalak) exportáljuk, mert ezen formátumokat ismerik az elemző szoftverek. Célszerű lehet a referencia-kezelő programok használata (EndNote, Zotero stb.) használni, így elkerülhető például a duplikáció, áttekinthetőbb, és hivatkozni is egyszerűbb ezek segítségével (Passas, 2024).
4. *Adattisztítás és előkészítés:* Az összegyűjtött, nyers adatbázist át kell vizsgálni és eltávolítani belőle a duplikációkat, ami főleg a több adatbázisból történő importálás okán fordulhat elő (Passas, 2024). Ezenkívül a szerzők nevét is érdemes átnézni, hiszen ugyanazon szerző többféleképpen is szerepelhet – vezetéknév, keresztnév variációk vagy elgépelés okán – vagy éppen hogy hasonló nevek nem kerülnek elválasztásra egymástól (Sanyal et al., 2019), azaz egységesítsük a metaadatokat minden szempontból. Ez végezhető Excellel, Pythonnal vagy R-rel is. A végeleges adatállománynak pontosnak és megbízhatónak kell lennie (Passas, 2024). A tisztítás folyamatát dokumentálni kell, és a módszertan kapcsán elérhetővé tenni, ha szükséges (pl. ha a bíráló kérdezi, vagy egyébként is beépítjük az Anyag és módszertan fejezetbe).
5. *Technika kiválasztása és alkalmazása:* a megtisztított adatbázisunk vizsgálatára használhatunk teljesítmény-, illetve hatáselemző technikát vagy tudományos hálózatok és kapcsolatok feltárására tudományterképészeti elemzést is (Passas, 2024). Ezeknek több fajtája létezik, pl. idézettségi elemzés (citation analysis), ko-citációs hálózatelemzés (milyen cikkeket emlegetnek sűrűn együtt, hiszen akkor vélhetően szoros kapcsolat van tartalmuk között), bibliográfiai kapcsolatháló (bibliographic coupling – ahol a tanulmányokban hivatkozott művek közötti hasonlóságot vizsgáljuk), kulcsszó- és tematikus elemzést (co-word analysis and thematic mapping – hogy milyen kulcsszó hálózatok fordulnak elő együtt gyakran, és kiderülhet belőle melyek például a marginális témák vagy feljövőben lévők) (Yu – Munoz-Justicia, 2020). Ezek a technikák önmagukban ritkábban használtak, inkább jellemző, hogy kombináltan jelennek meg.

Például egy átfogó bibliometriai elemzésben először az idézettségi elemzéssel a teljesítménymutatókat vizsgálhatjuk, utána a ko-citációs vagy kapcsolathálót elemezve a struktúrákat tárhatjuk fel, majd pedig a kulcsszóelemzéssel a tartalmi témákat azonosíthatjuk, így létrehozva egy átfogó tanulmányt az adott kérdésről. Ebben segít a szoftverek közül a VOSviewer vagy a CiteSpace, az R Bibliometrix csomagja (Passa, 2024).

7. *Eredményvizualizálás és értelmezés:* Ekkora adatmennyiségből készült eredményhalmaz könnyen értelmezhető vizuális formában jóval áttekinthetőbbé válik az olvasók számára. Ehhez használhatunk a szoftverek által generált gráfokat vagy mátrixokat, melyek a klasztereket, csoportokat, kapcsolatokat vagy a híd-szereplőket (összekötők két csoport között) mutatják be (Abramo et al., 2012; Passas, 2024). Az eredmények értelmezése csak akkor lehet sikeres, ha a kutató tisztában van az alaptémával, hiszen a szakmai kontextus szükséges a klaszterek, kapcsolatok pontos definiálásához, illetve a kapott eredményekből kiemelkedő publikációkat is át kell nézni és a feltárt kapcsolatot értelmezni kutatói szinten is (a szoftveresen túl). Használhatunk diagramokat, táblázatokat, kapcsolati irányokat jelző vizuális jeleket stb.

Marzi és kutatótársai (2024) javasolják, hogy az elkészült elemzés validálására szakértői interjú is alkalmazható triangulálás címén, így csökkenthető a torzító tényezők száma, és növelhető a hitelesség.

### **5. Tudományos cikkek tartalomelemzésének módszertana**

A tartalomelemzések sok esetben a tudományos cikkek megfelelő kvalitatív vagy kvantitatív feldolgozását, értékelését támogatják strukturált módon. Mivel a hallgatói tudományos munkák esetében ennek még az általános dokumentum- és tartalomelemzésnél lényegesebb szerepe lehet a szakirodalmi áttekintés vagy egy tudományos elmélet primer vizsgálata kapcsán, így ezt külön is tárgyaljuk. A tudományos cikkek tartalomelemzése segít a cikkekben fellelhető információk rendszerszintű feldolgozásában, bemutatásában. A vizsgálat megkezdése előtt fontos, hogy a konkrét témát pontosan meghatározzuk, a feltárás célját kijelöljük. Nagy segítség lehet nem csak az új (mint az ESG) és fejlődő irányzatok (mint az etnocentrizmus) feltárásban, de a különböző vélemények ütköztetésében vagy a különböző eredmények összehasonlító elemzésében. Ennek lépései a következők:

#### *1. Megfelelő cikkek kiválasztása és kódolása*

Ebben a szakaszban a korábbi kvalitatív eljárásokhoz hasonlóan meg kell határoznunk a kiválasztási kritériumokat, indikátorokat, melyek a kutató előzetes tapasztalata és a kutatási kérdése kapcsán irányulhatnak a cikkek relevanciájára, a módszertanra, a kutatási kérdéshez való illeszkedésre. Ezután meg kell határozni a kódolási keretrendszert, azaz például a cikkeket hogyan soroljuk be vagy osztályozzuk kutatási téma, módszertan vagy fontosabb következtetések, főbb megállapítások szerint. Ehhez a folyamathoz a legismertebb szoftverek – ha inkább kvalitatív elemzést készítünk – az NVivo, a MAXQDA vagy az Atlas.ti, illetve az R és a Python is javasolt szövegbányászathoz.

## *2. Kvalitatív tartalomelemzés lefolytatása*

Ebben a szakaszban a kiválasztott módszer segítségével végezzük el a cikkek főbb témáinak, elméleti modelljeinek vagy eredményeinek elemzését. Akkor érdemes kvalitatív eszközöket beemlíteni, amennyiben a szöveg mögötti üzeneteket, kontextust vagy szocio-ökonómiai környezetet, vonatkozásokat szeretnénk feltárni. A módszerek közül itt a legnépszerűbb a tematikus elemzés, mely segít a leggyakoribb témák és minták fel tárásában, illetve a narratív szintézis, ami a választott kutatási terület fejlődésének irányait képes feltárni, bemutatni. Példa: 50 releváns tudományos cikk elemzése az NVivo szoftver segítségével Fenntartható élelmiszermarketing (angol kulcsszó használat mellett) segített feltárni, hogy három főbb kutatási irány van jelen a kiválasztott cikkekben: a fenntarthatósággal kapcsolatos fogyasztói észlelések; a marketingkommunikáció hatékonyságának mérése a fenntartható élelmiszerek esetében, illetve az ökológiai címkék és minősítések fontossága a vásárlói döntéshozatalban. A cikkek kiválasztása történhet előzetes bibliometriai kvantitatív méréssel, melyet az előző fejezetben tárgyalunk.

## *3. Kvantitatív tartalomelemzés elvégzése*

Amennyiben szeretnénk számszerűsíthető, kvantitatív módon is vizsgálni a tartalmakat, végezhetünk szövegbányászatot és statisztikai elemzést, melyben a korábban említett Python használható, amivel az alábbi módszereket használhatjuk:

- Kifejezés/szó gyakoriságelemzése (word frequency analysis), mellyel feltárható, melyek a leggyakoribb kifejezések, kapcsolódó definíciók, a téma kapcsán feltárva a téma kulcsfogalmait, és kizárva olyan szavakat, mint az és, de, az stb. Ehhez a Pythonon belül alkalmazható a Natural Language Toolkit (NLTK) vagy a spaCy, majd ezek után ezek grafikusán is ábrázolhatók például szófelhőben például a Word Cloud használatával.
- Érzelmiltétel- vagy szentimentelemzés (sentiment analysis) (Hinek, 2024; Hinek – Papp-Váry, 2024), mely a szövegek érzelmi hangulatát (pozitív, semleges, negatív) igyekszik felismerni, megerősítve, hogy az adott cikk a témával kapcsolatban inkább támogató vagy kritikus hangvételi. Ehhez már bonyolultabb, gépi tanulást igénylő módszerek alkalmazhatóak, mint a VADER vagy a TextBlob, és végezhető manuális módon is, bár az automatikus a nagy mennyiségű adatok esetében gyorsabb lehet.
- A témamodellzés (topic modeling, LDA) (Hinek, 2024; Hinek – Papp-Váry, 2024) a szövegeket statisztikai modellel, valószínűségi számítás segítségével elemzi. Az LDA egy úgynevezett nem felügyelt (nem a kutató által betanított) gépi tanulási algoritmus, hanem önállóan, a témákat látens információként azonosítja a szövegekben. Hinek és Papp-Váry (2024) szerint a módszer jósága abból fakad, hogy nem előre meghatározott szavak vagy definíciók mentén keres, hanem a teljes dokumentumállományt vizsgálva keresi meg a témákat, hozzárendeli a jellemző kifejezéseket, és statisztikai számítások mentén alakít ki mintázatokat a témák szókinccse kapcsán. A megfelelő statisztikai tudás mellett alkalmazható dokumentumok, közösségimédia-tartalmak vagy akár más szöveges állományok szisztematikus értékelésére, főleg feltáró jelleggel, mivel a lényege az, hogy ne a kutató válasszon ki előzetes kulcsszavakat, hanem az algoritmus ismerje fel a főbb fogalmakat, kifejezéseket. Például segíthet beazonosítani, hogy

melyek azok a kifejezések, szavak, melyek már ismert kutatási témák esetében egyre gyakrabban kerülnek elő új kutatási irányokat előrevetítve.

- Időbeni trendek azonosítása kvantitatív mérésel, melynek elsődleges funkciója, hogy kiderüljön, egy adott téma mikor lett népszerű vagy veszítette el a tudomány érdeklődését.

Példaként említhető a kvantitatív tartalomelemzésre tudományos munkában, ha egy 300 cikkből álló mintán elemzzük a fenntartható fogyasztás (sustainable consumption) témakörét a Python és annak NLT könyvtárával, és a leggyakoribb kifejezések között megjelent az etikus fogyasztás, a karbon lábnyom és a greenwashing, vagy a témamodellezéssel az algoritmus 4 fő témába klaszterezi be a kapcsolódó irodalmakat.

A különböző módszerek kombinációjával mélyebb betekintést kaphatunk a kutatási területek fejlődésébe, alakulásába és a főbb kutatási kérdésekbe. Azt, hogy melyik az ideális az adatok mennyisége és a kipróbált eszközök használatának ismerete is befolyásolhatja.

### ***Irodalomjegyzék***

1. Abramo, G. – D'Angelo, C. A. – Di Costa, F. (2012): Identifying interdisciplinarity through the disciplinary classification of coauthors of scientific publications. *Journal of American Social Informatic Science Technology*, 63, 2206–2222. DOI: 10.1002/asi.22647
2. Bowen, G. (2009): Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. DOI: 10.3316/QRJ0902027
3. Cabezas-Clavijo, A. – Torres-Salinas, D. (2021): Bibliometric Reports for Institutions: Best Practices in a Responsible Metrics Scenario. *Frontiers in Research Metrics and Analytics*, 6, 696470. DOI: 10.3389/frma.2021.696470
4. Cserné Adermann, G. – Ponyi L. (2018): Kultúrakutatás módszertani alapjai. Tudástár X. kötet. Lakitelek: Nemzeti Művelődési Intézet.
5. Hinek M. (2024): Témák és érzelmek: Szállodai vendégvélemények vizsgálata témamodellezéssel és szentimentelemzéssel. In: A helyek arcvonásai (Szerk.: Michalkó G. – Rátz T. – Donka A.). Székesfehérvár, Budapest: Kodolányi János Egyetem, HUN-REN CSFK Földrajztudományi Intézet, Magyar Földrajzi Társaság, pp. 40–52.
6. Hinek, M. – Papp-Váry, Árpád (2024): Topic modelling of festival-visitors' opinions: A case study of Sziget Festival. *European Journal of Tourism Research*, 37, 3703. DOI: 10.54055/ejtr.v37i.3037
7. Jakusné Harnos É. (2023): A dokumentumelemzés módszertana a hadtudományi tanulmányok készítéséhez. *Hadtudomány*, 33, 157–173. DOI: 10.17047/Hadtud.2023.33.E.157
8. Marzi, G. – Balzano, M. – Caputo, A. – Pellegrini, M. M. (2025): Guidelines for Bibliometric-Systematic Literature Reviews: 10 steps to combine analysis, synthesis and theory development. *International Journal of Management Reviews*, 27, 81–103. DOI: 10.1111/ijmr.12381
9. Passas, I. (2024): Bibliometric Analysis: The Main Steps. *Encyclopedia*, 4(2), 1014–1025. DOI: 10.3390/encyclopedia4020065
10. Project Management Institute (PMI) (2020): *Projektmenedzsment útmutató*. 6. kiadás. Budapest: Akadémiai Kiadó. DOI: 10.1556/9789634545019

11. Sanyal, D. K. – Bhowmick, P. K. – Das, P. P. (2021): A review of author name disambiguation techniques for the PubMed bibliographic database. *Journal of Information Science*, 47(2), 227–254. DOI: 10.1177/0165551519888605
- Yu, J. – Muñoz-Justicia, J. (2020): A Bibliometric Overview of Twitter-Related Studies Indexed in Web of Science. *Future Internet*, 12(5), 91. DOI: 10.3390/fi12050091