



Magyar Agrár- és Élettudományi Egyetem

Szent István Campus

Műszaki Intézet

Adattechnológus-adatelemző szakmérnök

szakirányú továbbképzési szak

**Bibliográfiai adatok elemzésének összehasonlító vizsgálata
különböző adatvizualizációs szoftverekkel**

Belső konzulens: Lágymányosi Attila
egyetemi adjunktus

**Belső konzulens
intézete/tanszéke:** **Műszaki Intézet**

Külső konzulens: Záborszki Norbert
adatelemző

Készítette: **Dr. Szalay Zsigmond Gábor**

Gödöllő

2024

1 Tartalomjegyzék

2	Bevezetés és célkitűzések.....	4
2.1	Bevezetés.....	4
2.2	Célkitűzések	4
3	Szakirodalmi áttekintés	6
3.1	Bibliometria és a tudományterképezés.....	6
3.1.1	Bibliometria	6
3.1.2	Tudományterképezés	7
3.1.3	Kognitív térképek.....	9
3.2	Tudományos adatbázisok.....	10
3.2.1	Google Scholar.....	10
3.2.2	Web of Science	11
4	Anyag és módszertan	13
4.1	A Web of Science használata tudományos kutatásokhoz	13
4.1.1	A Web of Science platform.....	13
4.1.2	A Web of Science használatának lépései.....	13
4.1.3	A Web of Science részletes keresőjének használata	14
4.1.4	A részletes kereső használatának lépései	14
4.1.5	A Boole-i operátorok használata	14
4.2	Adatelőkészítés.....	15
4.2.1	Adatgyűjtés a Web of Science-ről	15
4.2.2	TXT fájlba exportálása Web of Science-ről.....	16
4.2.3	Adatkezelés, a munkafájl előállítása	17
4.3	Adatelemzés	18
4.3.1	A VOSviewer alkalmazása és jelentősége a tudományos kutatásokban.....	18
4.3.2	Adatbetöltés a VOSviewer-be	19
4.3.3	Adatelemzés VOSviewer-rel	20
4.3.4	Adatelemzés CitNetExplorer-rel.....	28
4.3.5	Adatelemzés Bibliometrix-szel	31
5	Eredmények.....	40
6	Következtetések és javaslatok.....	42
7	Összefoglalás	43
8	Irodalomjegyzék	44
9	Köszönetnyilvánítás.....	46
10	Mellékletek.....	47

Ábrajegyzék

1. ábra: Google Scholar logója.....	10
2. ábra: Web of Science logója	11
3. ábra: A Web of Science részletes keresőfelülete (advanced search).....	15
4. ábra: A Web of Science találati felülete	16
5. ábra. A Notepad++ logója.....	17
6. ábra: VOSviewer logója	18
7. ábra: A VOSviewer letöltési felülete.....	19
8. ábra: A kulcsszavak előfordulásának beállítása.....	20
9. ábra: A kulcsszavak megjelenési erősségének vizualizációja (Network Visualization)	21
10. ábra: A kulcsszavak évenkénti megjelenési erősségeinek vizualizációja (Overlay Visualization) ...	22
11. ábra: JSON export saját stop word lista készítése érdekében	22
12. ábra: A JSON fájl MS Excelben editálásra megnyitva	23
13. ábra: MS Excel-ben azonosított 'label'	23
14. ábra: A thesaurus módosításához használt .txt fájl.....	23
15. ábra: A kulcsszavak számának csökkenése a VOSviewer beállítási felületén	24
16. ábra: A csökkentett hálózati ábra.....	24
17. ábra: A LinLog-gal kapott átláthatóbb ábra.....	25
18. ábra: A kapcsolati háló alapjainak beállításai.....	25
19. ábra: A szerzők kapcsolata	26
20. ábra: Az országok kapcsolata	27
21. ábra: Hivatkozási háló	27
22. ábra: CitNetExplorer logója	28
23. ábra: A CitNetExplorer letöltési felülete	29
24. ábra: A CitNetExplorer adatfájl bekérő felülete.....	29
25. ábra: Hivatkozási kapcsolatok vizualizációja a különböző szerzők között	30
26. ábra: A CitNetExplorerben azonosított klaszterek megjelenése.....	30
27. ábra: A meghatározó irodalom elektronikus elérésének módja.....	31
28. ábra: Az alap forrásként azonosított cikk elektronikus megjelenése.....	31
29. ábra: A Bibliometrix logója	31
30. ábra: Az RStudio logója.....	32
31. ábra: A 'biblioshiny' parancs kiadása	33
32. ábra: A szakirodalmi információs fájl beolvasása Bibliometrix-be.....	33
33. ábra: A Bibliometrixbe belovastott adatok minősége	34
34. ábra: A legfontosabb adatok kinyerése bibliometrix-ben.....	35
35. ábra: Az évenkénti publikációk mennyisége a bibliometrix-ben.....	35
36. ábra: Ábra export a bibliometrix-ből.....	36
37. ábra: Háromdimenziós bibliográfiai összefüggések a Bibliometrix-ben	36
38. ábra: A legtöbb cikket felmutató források	37
39. ábra: A legtöbb hivatkozást felmutató források	37
40. ábra: Tematikus térkép a Bibliometrixben	38
41. ábra: Dendrogram a Bibliometrixben	39

2 Bevezetés és célkitűzések

2.1 Bevezetés

Az információk korában, ahol az adatok mennyisége és hozzáférhetősége rohamosan növekszik, kritikus jelentőségű az adatok hatékony elemzése és értelmezése. A tudományos és akadémiai közösség számára különösen fontos ez a feladat, hiszen a naprakész és releváns szakirodalommal való megismerkedés, valamint az adatok szisztematikus kiértékelése elengedhetetlen a kutatási projektek sikeréhez.

Ebben az összefüggésben a bibliográfiai adatok elemzése jelentős szerepet tölt be, és ezen belül is kiemelkedő jelentőséggel bírnak a speciális szoftverek által nyújtott lehetőségek. Ezek a szoftverek nem csupán segítik a kutatókat az adatok gyors és hatékony feldolgozásában, hanem lehetővé teszik számukra, hogy átfogóbb és részletesebb ismereteket szerezzenek az adott kutatási terület szakirodalmi helyzetéről és trendjeiről.

A diplomadolgozatom célja, hogy feltárja és elemezze ezeknek a bibliográfiai elemző szoftvereknek a jelentőségét és alkalmazhatóságát a tudományos kutatásban. Vizsgálni fogom ezen szoftvereknek a funkcionalitását, hatékonyságát és megbízhatóságát, valamint bemutatom azok előnyeit és korlátait. Emellett kitérek arra is, hogy hogyan járulnak hozzá ezek a szoftverek a tudományos ismeretek előrehaladásához, és milyen potenciált rejt magában az ilyen típusú technológiai fejlesztések további kutatások szempontjából.

A következő fejezetekben áttekintem a bibliográfiai elemző szoftverek legfontosabb jellemzőit, bemutatom azok működési elvét és alkalmazási területeit, valamint kritikus szemmel vizsgálom azok előnyeit és korlátait. Reményeim szerint ez a dolgozat hozzájárul majd kutató kollégáim információelemzés terén folytatott tevékenységének hatékonyabb és eredményesebb végrehajtásához.

2.2 Célkitűzések

1. **A bibliográfiai elemző szoftverek hatékonyságának vizsgálata:** Cél a különböző bibliográfiai elemző szoftverek hatékonyságának összehasonlító elemzése. Ennek során megvizsgálom, hogy milyen mértékben képesek ezek a szoftverek segíteni az adatok elemzésében és a szakirodalom feltárásában.

2. **A kutatói munkafolyamat hatékonyságának javítása:** Cél a bibliográfiai elemző szoftverek alkalmazásával a kutatói munkafolyamat hatékonyságának javítása. Ennek során megvizsgálom, hogy ezek a szoftverek mennyire képesek megkönnyíteni a kutatói tevékenységeket, és hogyan járulnak hozzá az eredményesebb kutatási eredmények eléréséhez.
3. **Az informatikai fejlesztések és innovációk vizsgálata:** Cél az informatikai fejlesztések és innovációk vizsgálata a bibliográfiai adatok elemzésében. Ennek során megvizsgálom, hogy milyen új technológiák és megközelítések jelentek meg ezen a területen, és hogyan befolyásolják ezek a fejlesztések a tudományos kutatást.

A dolgozat megírásával az alábbi rövidebb és hosszabb távú célokhoz szeretnék hozzájárulni:

- szoftveres bibliográfiai adatfeldolgozás bemutatása
- bibliográfiai adatok vizualizációja
- az egyes technológiai lehetőségek bemutatása, összehasonlítása
- komplex eszköztár összeállítása
- ajánlások megfogalmazása
- hibás alkalmazások bemutatása
- innovatív módszertani útmutató kialakítása

3 Szakirodalmi áttekintés

3.1 Bibliometria és a tudományterképezés

A bibliometria és a tudományterképezés olyan módszerek, amelyek segítségével elemzik és vizualizálják a tudományos publikációk jellemzőit és kapcsolatait. A bibliometria a tudományos publikációk és az azokban szereplő hivatkozások kvantitatív elemzésére összpontosít. Lehetővé teszi egy adott kutató vagy kutatási terület hatékonyságának és hatásának mérését. A tudományterképezés célja a tudományos ismeretek térbeli és tematikus ábrázolása, amelyeket gyakran vizuális eszközökkel, például hálózatok vagy térképek segítségével végeznek el. A két módszer összekapcsolása a tudományos kutatások szélesebb kontextusának megértésében és a kutatási trendek azonosításában is nagy segítséget nyújthat.

Ahogy az élet minden területén, úgy az akadémiai publikációk területén is rohamosan növekednek az elektronikus tartalmak száma, ami jelentősen megnehezíti a tudás naprakészen tartását, a meghatározó források ismeretét és a jelenleg futó trendek felismerését. A normál tapasztalati úton szerzett ismeretek immár töredezett kutatási ágakhoz vezetnek, így hiányos tudásmenedzsment alakulhat ki, ebből az okból kifolyólag a tudás felhalmozását és az evidenciák gyűjtését is megnehezítheti. Ennek eredményeképp a szakirodalomfelülvizsgálatok kulcsfontosságú szerepet töltenek be a múltbeli kutatási eredmények szintetizálásában és a tudásbázis hatékony jelenlegi felhasználásában. (Aria & Cuccurullo, 2017)

3.1.1 Bibliometria

Bibliográfiai adatokat a legkülönbözőbb hálózatokká lehet alakítani, majd ezekből származtatott hálózatokat lehet létrehozni. Az elemzésükhöz standard társadalmi hálózatelemzési módszereket lehet alkalmazni. Emellett a hivatkozási hálózatokat (és más WoS hálózatokat) időbeli hálózattá is alakíthatjuk, dinamikus jellemzőket tudunk benne megjeleníteni. (Batagelj & Cerinsek, 2013) Például a szlovén szerzőpáros által használt Pajek program is különböző szintű analíziseket tesz lehetővé, mellyel az adott egység halmaz felosztását és az osztályok zsugorítását végezhetjük el. Ez lehetővé teszi az érdeklődési területekre való fókuszálást és az adott célú hálózatcsökkentést.

Chaomei Chen cikkében bemutatott kibővített tudománytérképezési eljárás alkalmazható más területek elemzésére is. A kutatók ezeket a vizuális analitikai eszközöket felhasználhatják időszerű szakirodalmi áttekintések elvégzésére, valamint hatékonyabban találhatják meg a releváns, meghatározó publikációkat is. Ezek a kutatási eszközök egyre erőteljesebbek és elérhetőbbek lesznek. A meglévő technikák megnövelt körű alkalmazása pedig tovább szélesíti látókörünket, és segít az aktuális módszertani kihívások megértésében, melyeket le kell győznünk annak érdekében, hogy előre lépjünk a szakirodalmi hálózatok elemzésének fejlesztésében. (Chen, 2017)

3.1.2 Tudománytérképezés

A bibliometriai kutatásokban gyakran alkalmazzák a tudománytérképezést és a dokumentum csoportosítást annak érdekében, hogy leírják egy adott tudományos terület struktúráját és tartalmát. A tudománytérképezés általában a publikációk metaadatainak felhasználásával készíti el az adott terület hálózati reprezentációját, míg a dokumentum csoportosítás a dokumentumok szövegének elemzésével csoportosítja azokat hasonlóságuk alapján.

Az előbbi módszer főként arra összpontosít, hogy azonosítsa és vizualizálja a tudományos területek közötti kapcsolatokat, míg az utóbbi a dokumentumok tartalmi hasonlóságait keresi. Ezeknek a módszereknek az összehasonlítása lehetővé teszi számunkra, hogy betekintést nyerjünk a konkrét témákkal kapcsolatos nyelvi kifejezések és az azokat leíró tudományos struktúrák közötti kapcsolatba.

Az összehasonlítás segíthet azonosítani azokat a területeket is, ahol az elnevezések eltérnek vagy kontúrjaik homályosak, valamint, ahol eltérések tapasztalhatóak a két módszer által azonosított kulcsfogalmak között. Ezenfelül segíthet azonosítani azokat a területeket, ahol a két módszer kiegészíti egymást, ahol a kutatási módszertanok további finomításra szorulhatnak. Ennek köszönhetően az összehasonlítás egy fontos eszköz lehet a kutatási módszertanok fejlesztése és javítása terén. (Chen, et al., 2023)

A témák és témahálózatok vizualizálása a tudománytérképezés fontos része, melynek során a kulcsszavak és azok közötti kapcsolatok kerülnek előtérbe. Ezeket a csoportokat általában a témák megfelelőjeként értelmezzük, melyek két jellemzővel írhatóak le: a sűrűség és a centráliság. Ezeknek a témáknak a helyzetét megjeleníthetjük egy Stratégiai Diagramban, mely két tengely mentén ábrázolja a centráliság és a sűrűség rangsorértékeit.

Egy témát négyféle csoportba sorolhatunk a Stratégiai Diagramban való elhelyezkedésük alapján:

- A felső-jobb negyedben lévő témák jól fejlettek és fontosak a kutatási terület számára.
- A felső-bal negyedben lévő témák szintén jól fejlettek belső kapcsolatokkal, de kevésbé fontos külső kapcsolatokkal bírnak.
- Az alsó-bal negyedben lévő témák gyengén fejlettek és marginálisak.
- Az alsó-jobb negyedben lévő témák jelentik az alapokat, tekintve, hogy nagy az érdeklődés irántuk, de belső kapcsolataik nem fejlődnek. (Cobo, et al., 2010)

Dolgozatomban készítettem erre egy gyakorlati példát, melyet a 40. ábrán szemléltettem: Tematikus térkép a Bibliometrixben.

Cobo és társai (2011) a tudománytérképező szoftverek elemzését végezték el, melynek során kilenc különböző szoftvert vizsgáltak meg. Ezek a szoftverek különböző tulajdonságokkal rendelkeznek, de egyik sem jelölhető meg a „legjobb” jelzővel. A teljes tudománytérképezéshez egy elemzésen belül több szoftvert is használni kell, hogy összegyűjthessük az összes fontos információt, illetve a különböző nézőpontokat. Fontos megjegyezni, hogy a szoftverek különböző elemzési módszereket is kínálnak, amelyek hozzájárulnak a különböző ismeretek feltárásához. A vizualizációs megoldások eltérőek, így segítenek az eredmények értelmezésében és elemzésében. A különböző szoftverek együttműködése pozitív szinergiát eredményez, amely lehetővé teszi a rejtett tudás kiemelését az adatokból.

A tanulmány arra a következtetésre jutott, hogy bármely tudományterület alapos elemzése csak több szoftver kombinációjával végezhető el. Például, ha egy adott terület co-word elemzésére csak egy szoftver nyújt megnyugtató megoldást, ettől még más szoftverek is kínálhatnak fontos információkat: például absztraktok és címek kinyerése vagy időbeli változások követése. A különböző szoftverek kombinációja különböző nézőpontokat és információkat nyújthat, ami segít a teljes kép kialakításában.

Fontos megjegyezni, hogy a tanulmány nem fed le minden tudománytérképező szoftvert, sok kutató saját, egyedi szoftveres eszközöket és algoritmusokat használ. Ennek ellenére a tanulmány szerzői úgy vélik, hogy a különböző szoftverek együttműködése nagyobb mélységet és szélességet biztosít az elemzésekben. (Cobo, et al., 2011)

A bibliometriai térképezés és az információ-visszakeresés fontos eszközök a tudományos irodalom elemzésében. Ezek az eszközök főként a szavak, szókapcsolatok, idézetek és egyéb

jellemzők reprezentációin alapulnak. A standard bibliometriai modell a cikkek struktúrájából indul ki, amelyet különböző attribútumok, például idézetek és szavak, reprezentálnak. Az irodalom osztályozásának célja kettős: közvetlen osztályozás az elemzett cikkek alapján, vagy későbbi elemzések előkészítése. A hivatkozásokon alapuló térképezés alapjai már az 1960-as, 1970-es években kialakultak, és azóta folyamatosan fejlődnek különösen a szociális hálózatelemzés használatával.

Az irodalmi anyagok osztályozása és értelmezése számos módszer és eszköz felhasználását igényli, többek között a szöveges és hivatkozási elemek vizsgálatát, indexek alkalmazását. A kutatások folyamatos fejlesztése és az együttműködés a tudományos közösségek között elősegíti a szakirodalmi anyagok jobb megértését és értelmezését. (Glanzel, et al., 2019)

3.1.3 Kognitív térképek

Az informatikában gyakran használt adatszerkezetek közé tartoznak a gráfok vagy hálózatok. Ezeket gyakran diagramokkal jelenítik meg, hogy könnyebben értelmezhetővé váljanak. A gráfok automatikus rajzolásához – mint például állapotgráfok vagy adatáramlási gráfok – számos algoritmus létezik.

Az automatikus rajzolás területén számos kutatás történik, és egyre több algoritmus kerül fejlesztésre. Az egyik ilyen megközelítés, hogy a gráfban lévő csúcsok közötti távolságokat megfelelő geometriai távolságoknak tekintjük, és ennek alapján próbáljuk elhelyezni a csúcsokat a diagramon. Ennek az egyik módja, hogy egy virtuális dinamikus rendszert alkalmazunk, amelyben minden csúcsot rugók kötnek össze, és az optimális elrendezés az, amely minimalizálja a rendszer összes rugóenergiáját. (Kamada & Kawai, 1989)

A fuzzy kognitív térképek a rendszerelemek közötti kapcsolatrendszer ábrázolják gráf formájában, matematikai modellel, melyek érzékeltetik a rendszer elemek közötti kölcsönhatást. Az élek súlyozásával mutatják meg a kapcsolatokat, ahol a fuzzy súlyok a rendszer bizonytalanságát és jövőbeli viselkedését tükrözik. (Takács, 2014)

Qu és társai (2023) kutatása átfogóan elemzi a kognitív térképekkel kapcsolatos tanulmányokat 1970-től 2022-ig. Megállapítást nyert, hogy az elmúlt évtizedben exponenciálisan nőtt a kutatások száma ezen a területen, emellett a hivatkozások száma is pozitívan változott. A tanulmány rámutat arra, hogy a kutatások főként fejlett országokból származnak, továbbá a fejlődő országok részvétele még mindig alacsony. Az interdiszciplináris együttműködés és a magas impakt faktorú folyóiratokban való publikálás kulcsfontosságúak

voltak a terület fejlődésében. Összességében a kognitív térképek területe egyre nagyobb figyelmet kap, és ez a munka segíthet a kutatóknak átlátni a terület jelenlegi állását, meghatározni főbb irányait. (Qu, et al., 2023)

A kognitív térképek főleg a számítástechnikában, mérnöki területeken, földrajzban, neurológiában és pszichológiában koncentrálódnak. Az eredmények arra ösztönözik a kutatókat, hogy további, minőségi bibliográfiai tanulmányokat végezzenek a tudomány különböző területein, egyúttal jobb megoldásokat fejlesszenek ki az egyes területek számára. Az adatok elemzése egyéb előrejelző modellek fejlesztéséhez vezethet a bibliográfiai kapcsolatok általánosabb megértése érdekében. A kutatások eredményei hozzájárulhatnak a visszakeresési rendszerek fejlesztéséhez és a bonyolult bibliográfiai adatok kezeléséhez és értelmezéséhez. (Smiraglia & Leazer, 1999)

A bibliometrikus térképezés jellemzői:

- Hozzájárul a technológiai és a kapcsolódó tudományterületek vizualizációjához.
- Térképezésre több módszer is alkalmazható, mint például a co-citation analízis, bibliográfiai kapcsolat analízis, co-word analízis és co-class analízis.
- Az idősorokba rendezett térképek lehetővé teszik a tudásáramlás megismerését a tudomány és a technológia között, az országok vagy vállalatok között.
- Az idősoros térképeknek előjelző potenciáljuk is van: a feltörekvő vagy összefonódó technológiák korai felismerése. (van Raan, 2017)

3.2 Tudományos adatbázisok

3.2.1 Google Scholar

1. ábra: Google Scholar logója
(Forrás: www.scholar.google.com)



2004. novemberében a Google Scholar bevezetését hatalmas érdeklődés fogadta. A Google meglehetősen sikeres promóciós gépezete gyors népszerűséget hozott a Google Scholarnak, amelyet azóta is folyamatosan igyekeznek szinten tartani, illetőleg fokozni. A Google Scholar része a nagy Google keresőmotornak, mely teljes szöveges tudományos cikkeket, technológiai

jelentéseket, szakdolgozatokat és más dokumentumokat tartalmaz, valamint olyan weboldalakot, amelyeket 'tudományosnak' minősítenek. A Google állítása szerint teljes szöveges tartalom áll rendelkezésére minden nagyobb kiadótól, kivéve az Elseviert és az Amerikai Kémiai Társaságot, valamint olyan szolgáltatókat, mint a Highwire és az Ingenta. A Google Scholar indexének nagy része olyan teljes szöveges dokumentumokat tartalmaz, amelyeket mind kereskedelmi, mind nyílt forrású kiadók jelentettek meg. Emellett szakosodott bibliográfiai adatbázisokat, mint az OCLC Open WorldCat és a National Library of Medicine PubMed is indexál. A Google 2003. óta számos egyedi megállapodást kötött nagy kiadókkal a teljes szöveges tartalmak indexelésének lehetőségéért, amelyek másképpen nem lennének hozzáférhetők a nyílt weben. Bár a Google nem hozza nyilvánosságra a kiadók neveit vagy számát, akikkel megállapodást kötöttek, de könnyű belátni, hogy ez a szám igen magas lehet, hiszen minden kiadó érdekelt egy ilyen nagy láthatóságot biztosító globális vállalattal való együttműködésben. (Vine, 2006)

A Google Scholar rendkívül hasznos eszköz a kutatói munkát végző internetes közösség számára, hiszen lehetővé teszi, hogy gyorsan megtalálják a releváns tanulmányokat a világhálón elérhető milliárd oldal közül, sőt, sok esetben közvetlenül letölthessék ezek teljes szövegét. Fontos azonban ismerni a Google Scholar erősségei mellett a gyengeségeit is annak érdekében, hogy a kutatók eldönthessék, mikor és hogyan használják azt. A Google a Scholarral létrehozott egy nagy hasznos eszközt cikkek és egyéb szakirodalmak kereséséhez, ugyanakkor a Google Scholar-t már a kezdetek kezdetén alkalmatlannak tartották szakszerű és teljeskörű irodalomáttekintések készítésére a korlátozott keresési funkcionalitása miatt. (Giustin & Boulus, 2013)

3.2.2 Web of Science

2. ábra: Web of Science logója
(Forrás: www.webofknowledge.com)



A legnagyobb szakirodalmi adatbázis, amit használni szoktunk, az a Web Of Science, ami a WebofKnowledge portálon érhető el.

A Web of Science egy online tudományos információs platform, amelyet a Clarivate Analytics üzemeltet. Ez az egyik legismertebb és legbefolyásosabb tudományos adatbázis, amely

tartalmazza a világ számos tudományos területén megjelent cikkeket, folyóiratokat, konferenciakiadványokat, valamint az egymás közötti idézeteket és hivatkozásokat. A Web of Science jelentős szerepet játszik a kutatók, tudósok és más szakemberek számára a kutatási információk keresésében, az aktuális kutatási trendek nyomon követésében és a saját kutatási munkájukra való hivatkozások elemzésében.

A Web of Science által nyújtott szolgáltatások és funkciók közé tartozik:

1. Cikkek és kutatási eredmények keresése: Felhasználók keresést végezhetnek a tudományos publikációk széles körében, beleértve a cikkeket, konferenciakiadványokat és egyéb kutatási eredményeket.
2. Idézetek és hivatkozások: A Web of Science lehetővé teszi az idézetek és hivatkozások elemzését, amely segíthet a kutatóknak az adott téma iránti érdeklődés szintjének és a publikációk jelentőségének meghatározásában.
3. Témaközpontú csoportosítás: A platform lehetővé teszi a kutatási eredmények csoportosítását témakörök vagy tudományos területek szerint, ami segíthet a releváns információk megtalálásában.
4. Kutatói hálózatok és profilok: A Web of Science tartalmaz olyan funkciókat is, amelyek lehetővé teszik a kutatói hálózatok felderítését és a kutatók profiljainak megtekintését, ami segíthet a kapcsolatok kiépítésében és az együttműködések elősegítésében.
5. Keresési figyelmeztetések és egyéb szolgáltatások: A platform különféle keresési figyelmeztetéseket, hírleveleket és más testre szabott szolgáltatásokat is kínál, hogy a felhasználók mindig naprakészek legyenek az érdeklődési körükbe tartozó legújabb kutatási eredményekkel kapcsolatban.

A Web of Science portál használata előfizetéshez kötött, így a hozzáférési lehetőségek és a funkciók elérhetősége a felhasználók által választott előfizetés típusától függ.

4 Anyag és módszertan

4.1 A Web of Science használata tudományos kutatásokhoz

4.1.1 A Web of Science platform

A Web of Science (WoS) egy online tudományos információs platform, amely kiváló lehetőség a tudományos publikációkhoz és kutatási eredményekhez való hozzáféréshez. Ez a platform széles körű adatbázisokat kínál, melyek tartalmazzák a világ különböző tudományterületein megjelent cikkeket, konferenciakiadványokat és egyéb kutatási eredményeket. Ezenkívül lehetőséget nyújt az idézetek és hivatkozások elemzésére, valamint a kutatói hálózatok felderítésére és kapcsolatok kiépítésére a tudományos közösségben.

Kijelenthető, hogy a Nyugati Világban ez az egyik legelterjedtebb tudományos publikációs adatbázis, így megfelelő alapot nyújt a dolgozatomban kitűzött célok megvalósításához, a kiválasztott adatvizualizációs szoftverek teszteléséhez.

4.1.2 A Web of Science használatának lépései

1. *Hozzáférés és bejelentkezés:* A Web of Science platformhoz való hozzáférés általában intézményi előfizetéssel lehetséges, melyet az egyetemi vagy intézményi könyvtárak biztosítanak.
2. *Keresés:* A főoldalon található keresőmezőben lehet téma, cím vagy szerző szerint keresni. A részletes keresési lehetőségekhez a "Részletes keresés" opció választandó.
3. *Szűrés és finomítás:* A keresési eredmények megjelenése után szűrőkkel lehet a találati listát finomítani. Ezek közé tartozhatnak a dátumintervallumok, a tudományos területek, a folyóiratok, az idézetek száma stb.
4. *Eredmények böngészése:* A kapott eredmények szabadon böngészhetőek. Az érdeklődésre számot tartó találatokat könnyen elérhetjük, ha az már rendelkezik elektronikus példánnyal.
5. *Idézetek és hivatkozások:* A kiválasztott cikkek vagy kutatási eredmények részleteiben általában megtalálhatóak az idézetekre és hivatkozásokra vonatkozó adatok. Ezek segíthetnek abban, hogy felleljük milyen más munkák épülnek az adott kutatásra, és mennyire jelentős hatású az adott cikk.
6. *Profilok és kutatói hálózatok:* A Web of Science lehetőséget biztosít kutatói profilok megtekintésére és a kutatói hálózatok felderítésére. Ez segíthet abban, hogy megtaláljuk

azokat a kutatókat, akik hasonló témák iránt érdeklődnek, akik potenciálisan együttműködő partnereink lehetnek.

4.1.3 A Web of Science részletes keresőjének használata

A Web of Science részletes keresője lehetővé teszi a felhasználók számára a keresési eredmények szűrését, finomítását. Ez a funkció rendkívül hasznos tud lenni, ha specifikus keresési igény fogalmazódik meg egy tudományos projekt esetében.

4.1.4 A részletes kereső használatának lépései

1. Az *"Advanced search" (részletes kereső) lehetőség kiválasztása*: A főoldalon található keresőmező mellett általában található egy "Advanced search" link vagy gomb. Erre kattintva nyílik meg a részletes kereső felülete.

2. *Keresési feltételek megadása*: A részletes kereső felületén több mezőt találunk, amelyek lehetővé teszik a keresési feltételek pontosabb megadását. Ilyenek lehetnek például a cím, a szerző neve, a folyóirat neve, a kiadási évszám vagy igény szerinti kulcsszavak.

3. *Boole-i operátorok használata*: A részletes kereső lehetőséget biztosít a Boole-i operátorok: az "ÉS", "VAGY", "NEM" használatára a keresési lekérdezések finomításához. Ez segíthet abban, hogy kombináljuk a különböző keresési feltételeket a kívánt eredmény eléréséhez.

4. *Mezők kombinálása*: A részletes kereső lehetőséget biztosít különböző mezők kombinálására is. Például keresési lekérdezések adhatóak meg cím, szerző és kiadási év alapján is, hogy még pontosabb eredményeket kapjunk.

5. *Szűrés és finomítás*: Miután beállításra kerültek a keresési feltételek, a részletes kereső lehetővé teszi további szűrők és opciók alkalmazását a találatok további finomításához. Ilyenek lehetnek például a tudományos terület, a kiadvány típusa, az idézetek száma stb.

6. *Keresés indítása*: Amikor elkészültünk a keresési feltételek és a szűrők beállításával, a "Keresés" gombra kattintva a találatok megjelennek.

4.1.5 A Boole-i operátorok használata

1. *ÉS (AND) operátor*: Az "AND" operátor használatakor a keresés csak azokat az elemeket fogja visszaadni, amelyek minden feltételnek megfelelnek. Például: "digitalizáció AND mezőgazdaság" azt jelenti, hogy csak azok a cikkek jelennek meg, amelyekben mind a "digitalizáció", mind a "mezőgazdaság" szó előfordul.

2. *VAGY (OR) operátor*: A "VAGY" operátorral ellátott keresés minden elemet visszaad, amely legalább egy feltételnek megfelel. Például: "digitalizáció OR informatika" azt jelenti, hogy azok a cikkek is megjelennek, amelyekben vagy a "digitalizáció", vagy az "informatika" szó előfordul.
3. *NEM (NOT) operátor*: A "NEM" operátor kizárja azokat az elemeket, amelyek tartalmazzák a megadott kifejezést. Például: "digitalizáció NOT ipar" azt jelenti, hogy a cikkekben előforduló "digitalizáció" szavakat keresi, de kizárja azokat, amelyek tartalmazzák az "ipar" szót is.
4. *Zárójelek használata*: A zárójelek használata segíthet abban, hogy tisztább és összetettebb kereséseket végezzünk. Például: "(digitalizáció OR informatika) AND mezőgazdaság" azt jelenti, hogy olyan cikkeket keresünk, amelyekben vagy a "digitalizáció" vagy az "informatika" szavak előfordulnak, és tartalmazzák a "mezőgazdaság" szót is.

4.2 Adatelőkészítés

4.2.1 Adatgyűjtés a Web of Science-ről

Jelen esetben az alábbi keresés került definiálásra, az alábbi utasítást alkalmaztam:

$$TS=((„agri*” OR „agro*”) AND („artificial intelligence” OR „AI”))$$

3. ábra: A Web of Science részletes keresőfelülete (advanced search)
(Forrás: saját munka a Web of Science-en)

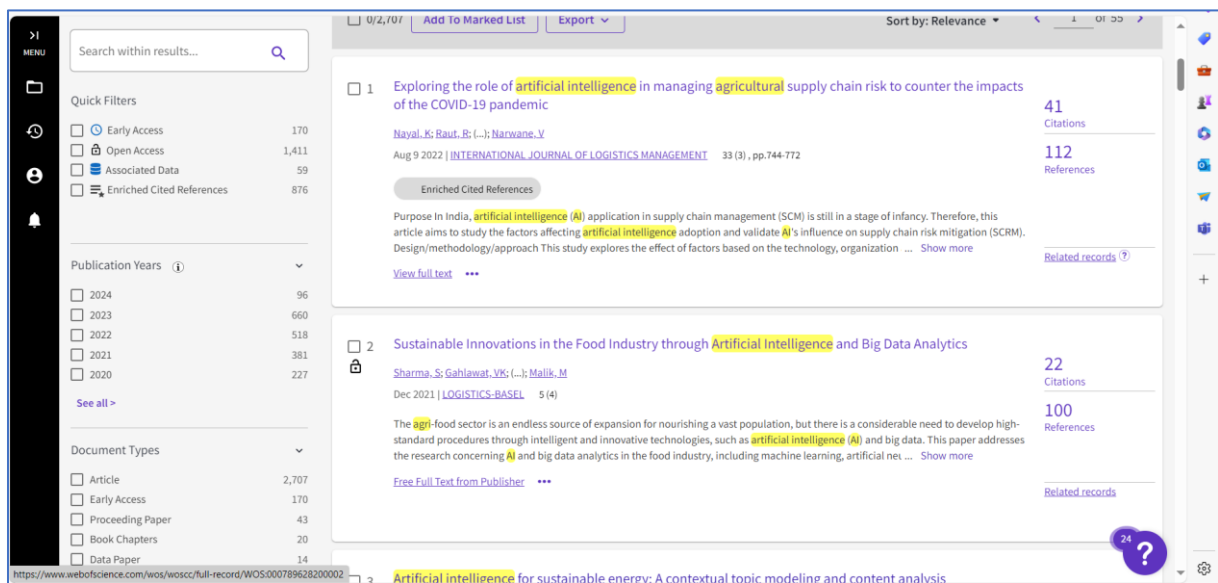
The screenshot shows the Clarivate Web of Science interface. At the top, there are navigation tabs for 'Egyetemi rendszerek', 'Oktatás', 'E-learning', 'Tudományos', 'Magán', and 'Szakmai'. The main header includes the Clarivate logo, language settings (English), and a 'Products' menu. The search bar contains the query: `TS=((„agri*” OR „agro*”) AND („artificial intelligence” OR „AI”))`. Below the search bar, there are buttons for 'Analyze Results', 'Citation Report', and 'Create Alert'. The search results section shows '2,707 results from Web of Science Core Collection for:'. A 'Refined By' section shows 'Document Types: Article' selected. The results list includes a single entry with the title 'Exploring the role of artificial intelligence in managing agricultural supply chain risk to counter the impacts of the COVID-19 pandemic' by Nayal, K; Raut, B; (-); Narwane, V. The entry has 41 citations and is from the journal 'INTERNATIONAL JOURNAL OF LOGISTICS MANAGEMENT', issue 33(3), pp.744-772, published in Aug 9 2022.

Az 'OR' operátorok használatával az azonos értelemben használt szavak kerültek be a keresési paraméterekbe. Hiszen angol nyelvterületen egyaránt találkozhatunk „agri” (pl.: agriculture) és „agro” (pl.: agroproduction) kezdetű kifejezésekkel is. A '*'-gal¹ definiáltam, hogy itt szóeleji egyezőségeket keresek. Az „AI” pedig az „artificial intelligence” rövidítése, így adekvát, hogy mindkettő kifejezés benne legyen a találati listában. A Web of Science részletes keresőjét és találati listáját az 3. ábra és 4. ábra szemlélteti.

A fenti keresésre 2 707 találatot kaptam 2024. március 16-án.

Érdeemes külön figyelmet fordítani a „Documnet Types” szűrési lehetőségre.

4. ábra: A Web of Science találati felülete
(Forrás: saját munka a Web of Science-en)



4.2.2 TXT fájlba exportálása Web of Science-ről

A Web of Science-ből TXT fájlba való exportálás lehetővé teszi a felhasználók számára, hogy a keresési eredményeket egy egyszerű szövegfájlba menthessék el. Ez a funkció szükséges lehet, ha a felhasználó keresési eredményeit más szövegszerkesztő vagy adatfeldolgozó szoftverben szeretné megnyitni, hogy további elemzéseket végezzen.

TXT fájl exportálásának folyamata Web of Science-ből:

1. *Keresési eredmények kiválasztása:* Először a Web of Science adatbázisban végzett keresés eredményéből ki kell jelölni az exportálandó adatok legfontosabb jellemzőit. Ez lehet például egy cikkek listája, konferencia-proceedingek vagy más kutatási eredmények.

¹ A '*' egy úgynevezett 'wildcard', azaz helyettesítő karakter.

2. *Exportálási lehetőség kiválasztása:* Miután kijelöltük a keresési eredményeket, keressük meg az „Export” lehetőséget a felületen.

3. *TXT fájl formátum kiválasztása:* A Web of Science lehetőséget biztosít különböző fájl formátumok kiválasztására. TXT fájlok esetében az úgynevezett „Plain text file” opció választandó.

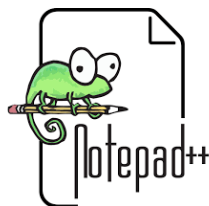
4. *Exportálás megerősítése:* Amennyiben elegendő a szerző, a cím és a szakirodalmi adatok (author, title, source), akkor 1000 rekordot lehet egyszerre letölteni, azaz 1000 cikk adatai kerülnek a mentésbe. Azonban ettől sokkal jobban elemezhető adatvagyon eredményez, ha a „Full Record and Cited References” opciót választjuk. Ebben az esetben 67 a cikkekre vonatkozó irodalmi adat lesz az adatbázisunkban, amelyet 68-ként követ az, hogy az adott cikk milyen irodalmakra hivatkozik. Ilyenkor 500 rekordban lesz maximalizálva az egyszerre letölthető sorok száma. Külön figyelmet kell arra fordítani, hogy az utolsó rekord számát pontosan kell megadni, egy bővebb 500-as tartomány hibára fog futni!

5. *Fájl letöltése:* A Web of Science el fogja kezdeni az exportálási folyamatot, és felajánlja a .txt fájl letöltésének lehetőségét. Ez általában a letöltési helyre irányító hivatkozás vagy gomb formájában jelenik meg.

4.2.3 Adatkezelés, a munkafájl előállítása

A további munkához célszerű a legegyszerűbb szövegszerkesztő használata. Window-os környezetben ez kézenfekvően a Notepad program. Ugyan lényegesen kisebb funkcionalitással bír, mint az MS Word, de nagyon megbízhatóan kezel több százezer, akár több millió adatsoros fájlokat is. A biztos adatkezeléshez némi kényelmi funkciót biztosít a Notepad++ szoftver, mely e tekintetben a legnépszerűbb alkalmazásnak számít. (5. ábra. A Notepad++ logója)

5. ábra. A Notepad++ logója
(Forrás: <https://notepad-plus-plus.org>)



Mivel a Notepad++ egyszerre több .txt fájl megnyitását is lehetővé teszi, így a lementett fájlok összemásolása nagyon könnyen megoldható. A Notepad++ másik kényelmi funkciója is

azonnal hasznosítható információval szolgál, nevezetesen megnézhető az így keletkezett fájl sorainak száma, melyet jelen esetben 302 851 sornak találtam.

Még egy nagyon fontos adatmódosításra van szükség! Az összesített fájlban az alábbi karakterlánc cserélendő:

FN Clarivate Analytics Web of Science

Helyette erre lesz szükség:

FN ISI Export Format VR.

Illetve ez üres karakterre cserélendő:

VR 1.0

Ezzel elkészült a szakirodalmi információs fájlom.

4.3 Adatelemzés

4.3.1 A VOSviewer alkalmazása és jelentősége a tudományos kutatásokban

6. ábra: VOSviewer logója

(Forrás: <https://www.vosviewer.com>)



A VOSviewer egy olyan szoftver, amely lehetővé teszi a tudományos adatok vizualizációját és elemzését, különösen a tudományos publikációk, szerzők és témák közötti kapcsolatok tekintetében. Ez a szoftver széles körben használt eszköz a tudományos kutatásokban, mivel segítségével könnyen áttekinthetőek a kutatási területek közötti kapcsolatok és folyamatok.

A VOSviewer alkalmazási területei az alábbiak szerint csoportosítható.

Hálózati térképek készítése: A VOSviewer lehetővé teszi hálózati térképek létrehozását a tudományos publikációk, szerzők vagy más entitások közötti kapcsolatok alapján. Ezek a térképek segíthetnek azonosítani a központi témákat és azokat a területeket, amelyeken a legtöbb kutatás zajlik.

Szerzői hálózatok elemzése: A szoftver segítségével vizualizálhatók és elemezhetők a szerzők közötti kapcsolatok, ami lehetővé teszi a központi szerzők, az együttműködések és más releváns jellemzők azonosítását.

Témakörök azonosítása: A VOSviewer segítségével könnyen azonosíthatók a központi témakörök és kulcsszavak egy adott tudományos területen vagy adathalmazon belül.

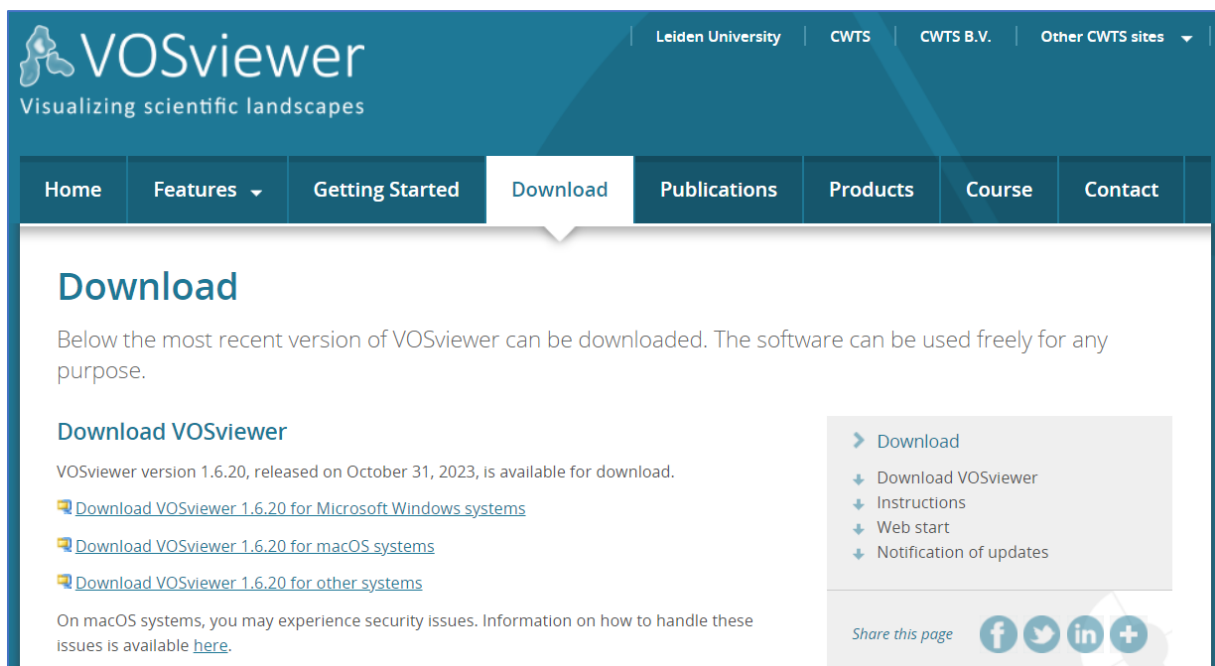
Hivatkozások és idézetek elemzése: A VOSviewer segítségével könnyen vizualizálhatók a hivatkozások és idézetek hálózatai a tudományos publikációk között, amely segíthet az adott területen a leginkább hivatkozott vagy befolyásos cikkek azonosításában.

A VOSviewer széles körben elterjedt eszköz a tudományos kutatásokban, mivel segítségével könnyen és hatékonyan vizualizálhatók és elemzhetők a tudományos cikkek metaadatai. A szoftver lehetővé teszi a kutatók számára, hogy mélyebb betekintést nyerjenek a kutatási területekbe, folyamatokba; és segítsen az új irányok és lehetőségek azonosításában.

4.3.2 Adatbetöltés a VOSviewer-be

A VOSviewer ingyenesen letölthető, melyet a 7. ábra szemléltet.

7. ábra: A VOSviewer letöltési felülete
(Forrás: <https://www.vosviewer.com/download>)



Mivel még nem rendelkezem elkészült hálózatokkal, így a „Create a map based on bibliographic data” opciót kell választanom, azaz alap bibliográfiai adatokból dolgozom, és tárok fel új összefüggéseket.

Forrásadataim a Web of Science-ről származnak, így a következő panelen a „Read data from bibliographic database files” opciót szükséges bejelölnöm.

Ezek után nyílik lehetőség az előzőleg legyártott szakirodalmi információk fájl betöltésére.

4.3.3 Adatelemzés VOSviewer-rel

A legegyszerűbb vizsgálat úgy végezhető, hogy megnézzük melyek azok a kifejezések, amelyek egyszerre fordulnak elő. Az egyszeri előfordulás alapján a program csinál egy meglehetősen nagy, úgynevezett szomszédossági mátrixot.

A szomszédossági mátrix egy olyan matematikai eszköz vagy adatszerkezet, amelyet gyakran használnak a hálózatkutatásban és a gráfelméletben. Egy adott hálózat vagy gráf esetében a szomszédossági mátrix megadja az egymással közvetlen kapcsolatban álló csúcsok vagy csomópontok közötti kapcsolatokat vagy élek jelenlétét esetleges hiányát.

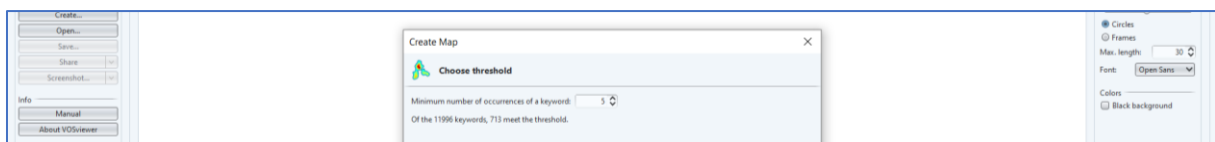
Egy szomszédossági mátrixban az oszlopok és a sorok általában a hálózatban szereplő csúcsokat vagy csomópontokat jelentik, és az egyes cellák értékei megadják azok közötti kapcsolatokat. Ha az adott csúcsok között van él a hálózatban, akkor a megfelelő cella értéke általában 1 vagy egy másik előre meghatározott érték, míg ha nincs kapcsolat, akkor az érték általában 0 vagy egy másik jelző érték lehet.

Angolul a szomszédossági mátrixra általában "adjacency matrix" kifejezést használják. Ez a kifejezés pontosan tükrözi a mátrix célját és tartalmát, vagyis hogy leírja az egyes csúcsok közötti kapcsolatokat vagy szomszédságot a hálózatban.

A kulcsszavak esetében a „Co-occurrence” és az „All keywords” opciók választását tekinthetjük alapbeállításnak. Ezzel a kulcsszavak bármilyen előfordulását validnak tekintjük.

A szoftver alapértéknek az 5-öt tekinti, azaz minimum 5 előfordulást elért kifejezést tekint 1-nek. Ez alapján a 11 996 kulcsszóból 713 esetről talált küszöbértéket elérő vagy azt meghaladó értéket. Természetesen ez a küszöbérték állítható: növelésével csökken a küszöbértéket meghaladók száma. (8. ábra)

8. ábra: A kulcsszavak előfordulásának beállítása
(Forrás: saját munka VOSviewer-ben)



Ez a 713-as érték már nem tartalmazza a szövegelemzésben úgynevezett 'stop word'-öket.

A stop word (magyarul „stop szó”) olyan szó vagy szókapocs, amelyet általában figyelmen kívül hagynak egy adott szövegfeldolgozási feladat során, mivel nem hordoznak jelentőséget vagy információt a szöveg tartalmával kapcsolatban. Ezek a szavak általában gyakori nyelvi elemek, például kötőszavak, névmások, igekötők stb.

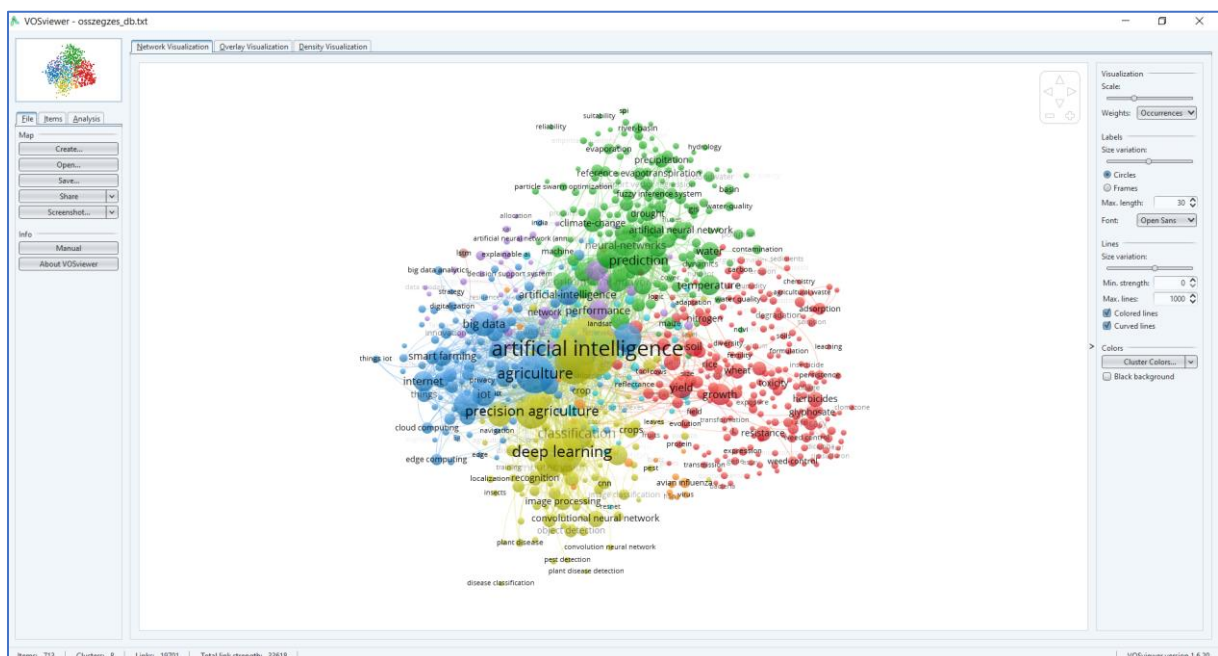
A stop szavak kiszűrése vagy eltávolítása előnyös lehet olyan feladatokban, mint például a keresési lekérdezések feldolgozása, a szövegbányászat, a dokumentumok címkézése vagy a természetes nyelvű feldolgozás (NLP), mivel a szövegek összehasonlítása vagy elemzése során ezek a szavak nem jelentenek fontos információt, és gyakran csak zajt jelentenek.

Példák angol 'stop word'-ökre: "the", "a", "an", "and", "of", "in", "on", "is", "are", "am", "be", "have", "has", "had", "was", "were" stb.

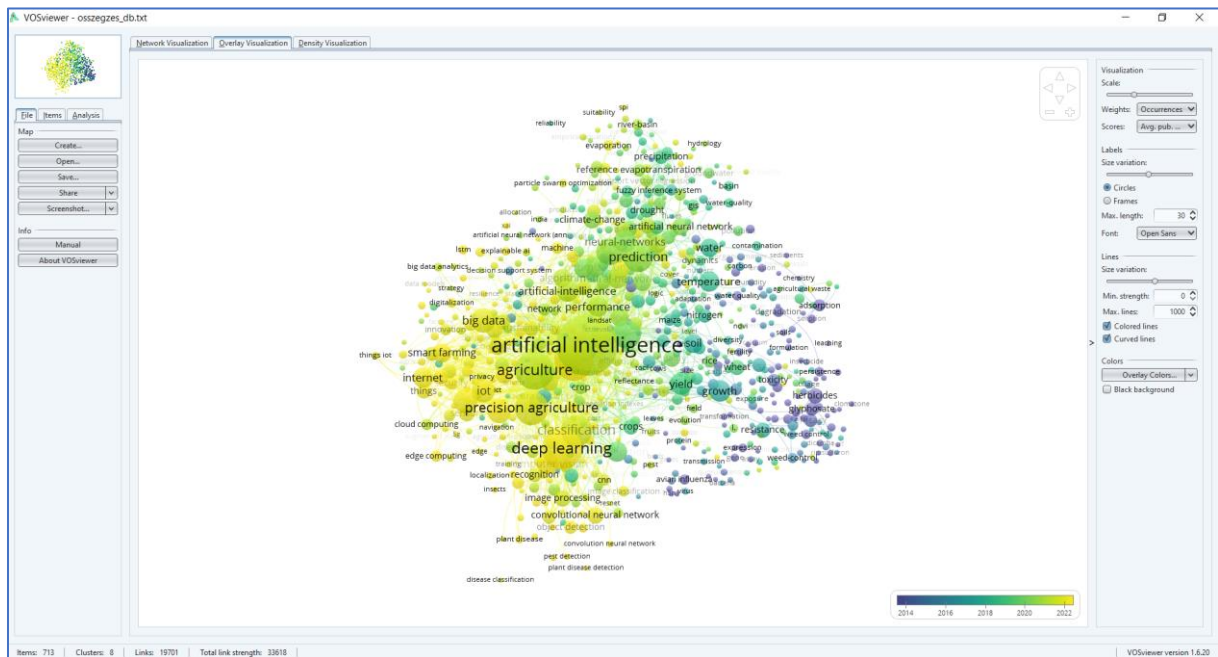
A stop szavakat gyakran egy előre definiált lista alapján határozzák meg, és ezt a listát az adott feladat vagy alkalmazás jellegétől függően állítják össze. A stop szavak eltávolítása segíthet abban, hogy a fontosabb vagy releváns információk kiemelkedjenek a szövegből, és hogy a gépi tanulási vagy adatfeldolgozási modellek hatékonyabban dolgozhassanak a szöveges adatokkal.

A 9. ábra mutatja a végeredményt. Az azonos színnel jelölt szavak egy klaszterbe tartoznak. Minél nagyobb egy szót jelölő kör, annál többször fordul elő a szakirodalmi információk fájlban. A 10. ábra pedig megjeleníti, hogy az egyes kulcsszavak az évek során milyen erővel jelentek meg a szakirodalomban.

9. ábra: A kulcsszavak megjelenési erősségének vizualizációja (Network Visualization)
(Forrás: saját munka VOSviewer-ben)



10. ábra: A kulcsszavak évenkénti megjelenési erősségeinek vizualizációja (Overlay Visualization) (Forrás: saját munka VOSviewer-ben)



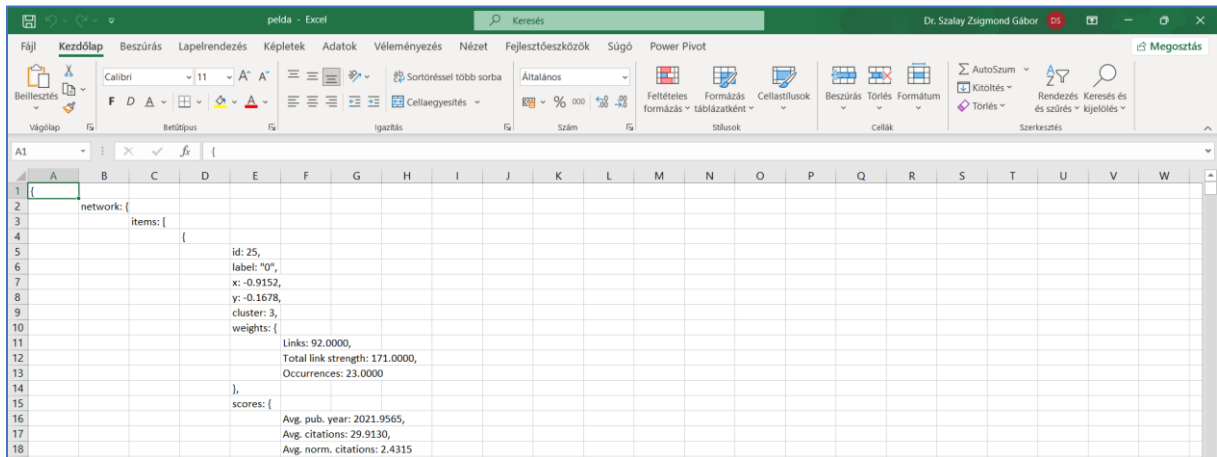
Minden tudományterületnek megvannak azok a hívószavai (buzzword), amelyek uralják a témát, de szövegelemzés szempontjából nem tartalmaznak semmilyen többletinformációt. Ezeket célszerű egy saját tezaurusszal eliminálni a kiértékelésből, hiszen evidenciákat jelenítenek meg.

Egyfelől a VOSviewer-ben is megnézhetjük a kulcsszavak top listáját, de szoftveren kívül nyílik igazán lehetőség statisztikai vizsgálatokra. Egyik legegyszerűbb megoldásnak kínálkozik például, ha a kulcsszavakat egy JSON fájlba exportáljuk (11. ábra), amit akár Notepad++ programmal, de akár MS Excel-lel is elemezhetünk (12. ábra, 13. ábra). Ennek során legyártható a kutató saját stop word listája.

11. ábra: JSON export saját stop word lista készítése érdekében (Forrás: saját munka)

```
1  "network": {
2    "items": [
3      {
4        "id": 25,
5        "label": "0",
6        "x": -0.9152,
7        "y": -0.1678,
8        "cluster": 3,
9        "weights": {
10         "links": 92.0000,
11         "total link strength": 171.0000,
12         "occurrences": 23.0000
13       },
14       "scores": {
15         "Avg. pub. year": 2021.9565,
16         "Avg. citations": 29.9130,
17         "Avg. norm. citations": 2.4315
18       }
19     ]
20   }
```

12. ábra: A JSON fájl MS Excelben editálásra megnyitva
(Forrás: saját munka)



13. ábra: MS Excel-ben azonosított 'label'
(Forrás: saját munka)

{				0
id: 94,				0
label: "absorption",			1	"absorption",
x: 0.8411,				0
y: -0.3397,				0
cluster: 1,				0
weights: {				0

A statisztikai eljárások mellett a kutató ember, a szakértő szerepe sem mellékes. A végső döntést saját szakmai felkészültsége, tapasztalatai alapján kell meghoznia.

Saját tezaurusz fájl létrehozására a Notepad a legalkalmasabb eszköz. A fájl struktúrája az alábbiak szerint néz ki:

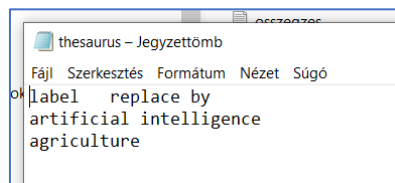
label 'TAB' replace by

artificial intelligence 'TAB' „semmi”

AI 'TAB' „semmi”

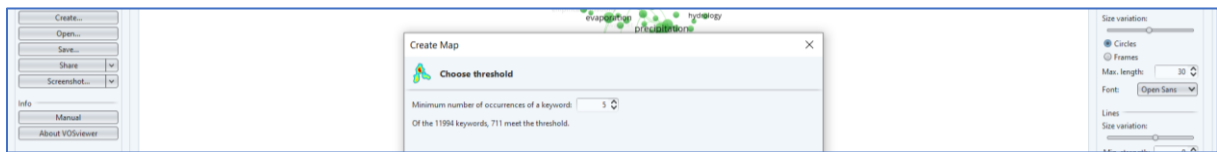
Ebben a példában az „artificial intelligence” és az „AI” kulcsszavak vettem ki a vizsgált halmazból. (14. ábra)

14. ábra: A thesaurus módosításához használt .txt fájl
(Forrás: saját munka Notepad-ben)



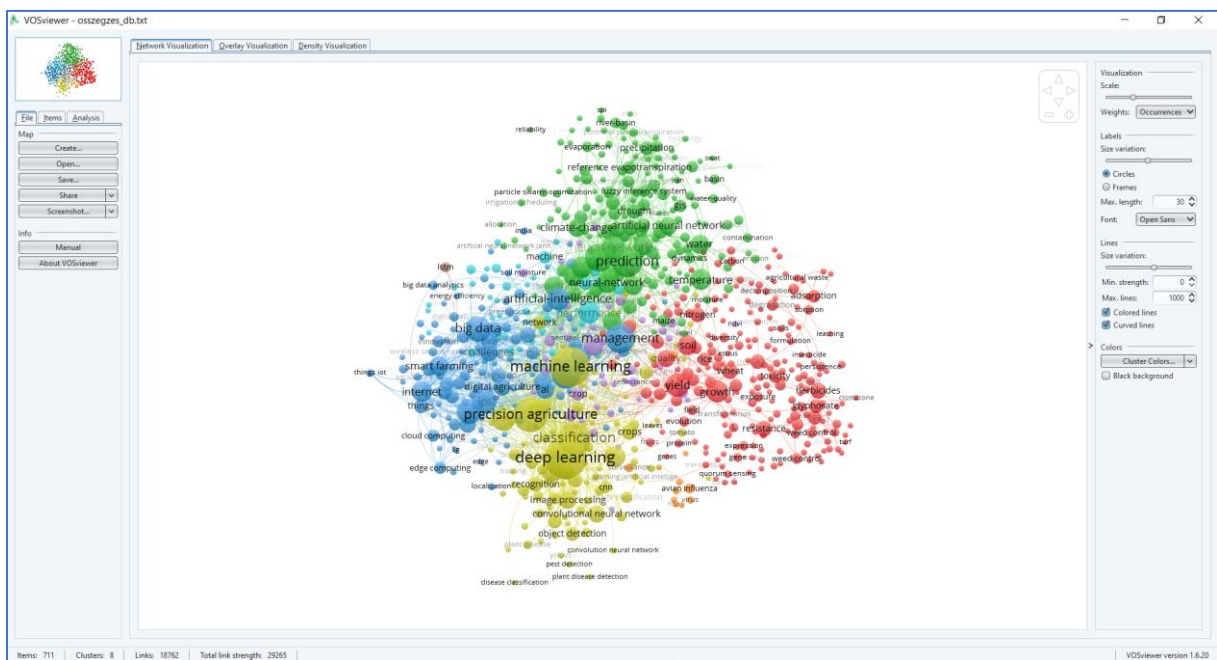
A .txt fájlban szereplő kettő kulcsszóval láthatóan 711-re csökkent is az elemek száma. (15. ábra)

15. ábra: A kulcsszavak számának csökkenése a VOSviewer beállítási felületén
(Forrás: saját munka VOSviewer-ben)



A VOSviewer-ben is elérhettem volna hasonló eredményt a kulcsszavak toplistáját mutató felugró ablak jelölőnégyzeteit használva. Azonban minden vizualizáció futtatásakor a program „elfelejti” ezeket a jelöléseket, azaz minden egyes alkalommal újra kell a jelölőnégyzeteket használni. Könnyen belátható, hogy ez milyen hibázási lehetőséghez vezet. További előnye a tezaurusz fájl használatának, hogy a későbbiekben is, egy másik vizsgálat során is felhasználható, így a kutató ki tudja alakítani a saját kutatási területén alkalmazandó tezaurusz fájlt. A 16. ábra mutatja a már csökkentett kulcsszó listával készült diagramot.

16. ábra: A csökkentett hálózati ábra
(Forrás: saját munka VOSviewer-ben)



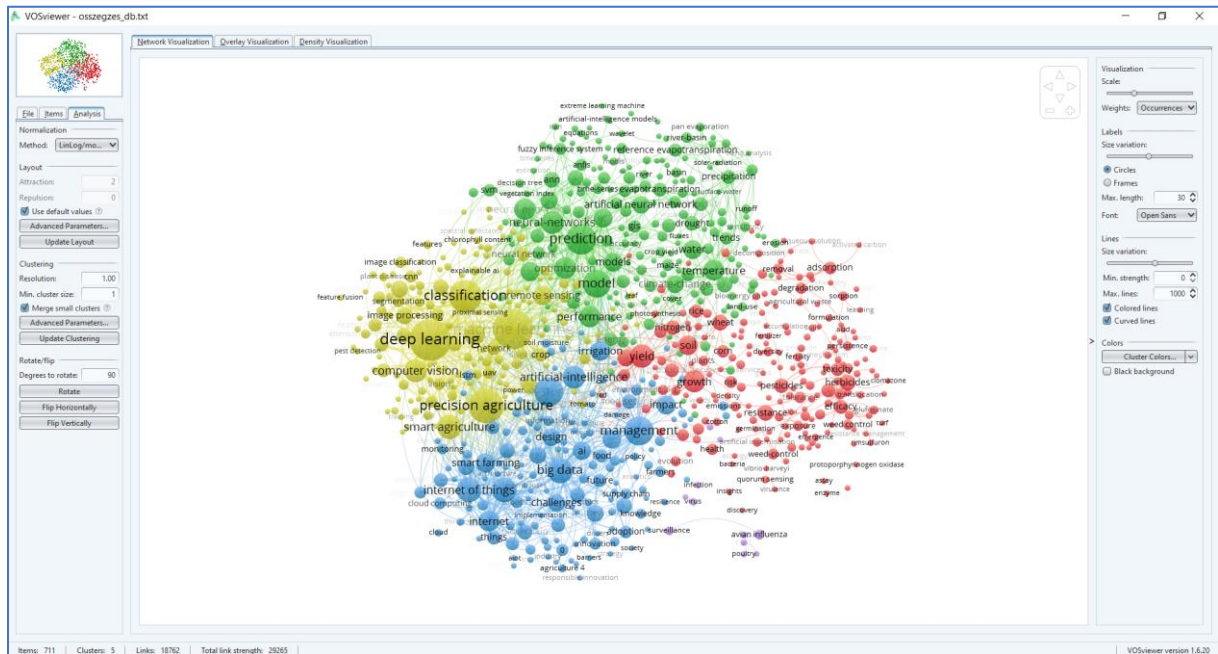
Az ábra áttekinthetőségére jótékony hatással lehet a loglineáris modularitás bekapcsolás, melynek eredményét a 17. ábra mutatja.

A VOSviewer számos típusú összefüggésre képes ábrákat készíteni a tudományos adatok vizualizációjához. (18. ábra) Ezek közül néhány a következő:

Egyidejű előfordulás: A co-occurrence elemzés célja az, hogy feltárja és meghatározza, hogy az adott elemek, kulcsszavak hogyan kapcsolódnak egymáshoz vagy hogyan jelennek meg

együtt bizonyos körülmények között. Ez lehetővé teszi az összefüggések, mintázatok vagy trendek azonosítását az adott adathalmazban. (17. ábra)

17. ábra: A LinLog-gal kapott átláthatóbb ábra
(Forrás: saját munka VOSviewer-ben)



18. ábra: A kapcsolati hálóknak alapjainak beállításai
(Forrás: saját munka VOSviewer-ben)

Create Map

Choose type of analysis and counting method

Type of analysis:	Unit of analysis:
<input type="radio"/> Co-authorship	<input checked="" type="radio"/> Documents
<input type="radio"/> Co-occurrence	<input type="radio"/> Sources
<input type="radio"/> Citation	<input type="radio"/> Authors
<input checked="" type="radio"/> Bibliographic coupling	<input type="radio"/> Organizations
<input type="radio"/> Co-citation	<input type="radio"/> Countries

Counting method:

Full counting

Fractional counting

A VOSviewer számos típusú összefüggésre képes ábrákat készíteni a tudományos adatok vizualizációjához. (18. ábra) Ezek közül néhány a következő:

Egyidejű előfordulás: A co-occurrence elemzés célja az, hogy feltárja és meghatározza, hogy az adott elemek, kulcsszavak hogyan kapcsolódnak egymáshoz vagy hogyan jelennek meg együtt bizonyos körülmények között. Ez lehetővé teszi az összefüggések, mintázatok vagy trendek azonosítását az adott adathalmazban. (17. ábra)

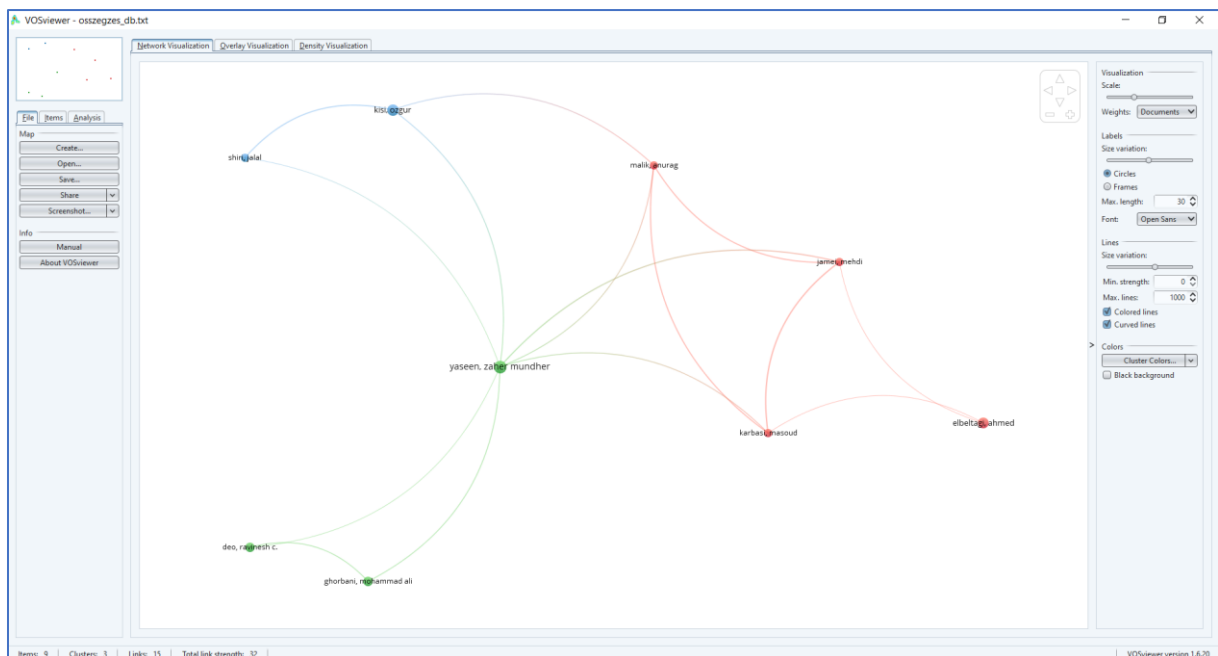
Szerzői kapcsolatok: Az egyik leggyakoribb típusú ábra a szerzői kapcsolatok ábrázolása, amely az együttműködések vagy közös munkák alapján jeleníti meg a szerzők közötti kapcsolatokat. (19. ábra)

Országok közötti kapcsolatok: A VOSviewer segítségével ábrázolhatóak az országok közötti kapcsolatok is, például a nemzetközi együttműködések vagy közös kutatási projektek. (20. ábra)

Hivatkozási kapcsolatok: Másik fontos típusú ábra a hivatkozási kapcsolatok ábrázolása, ami azt mutatja, hogy mely publikációk hivatkoznak más publikációkra, és ezzel az idézeti hálózatot jeleníti meg. (21. ábra)

19. ábra: A szerzők kapcsolata

(Forrás: saját munka VOSviewer-ben)

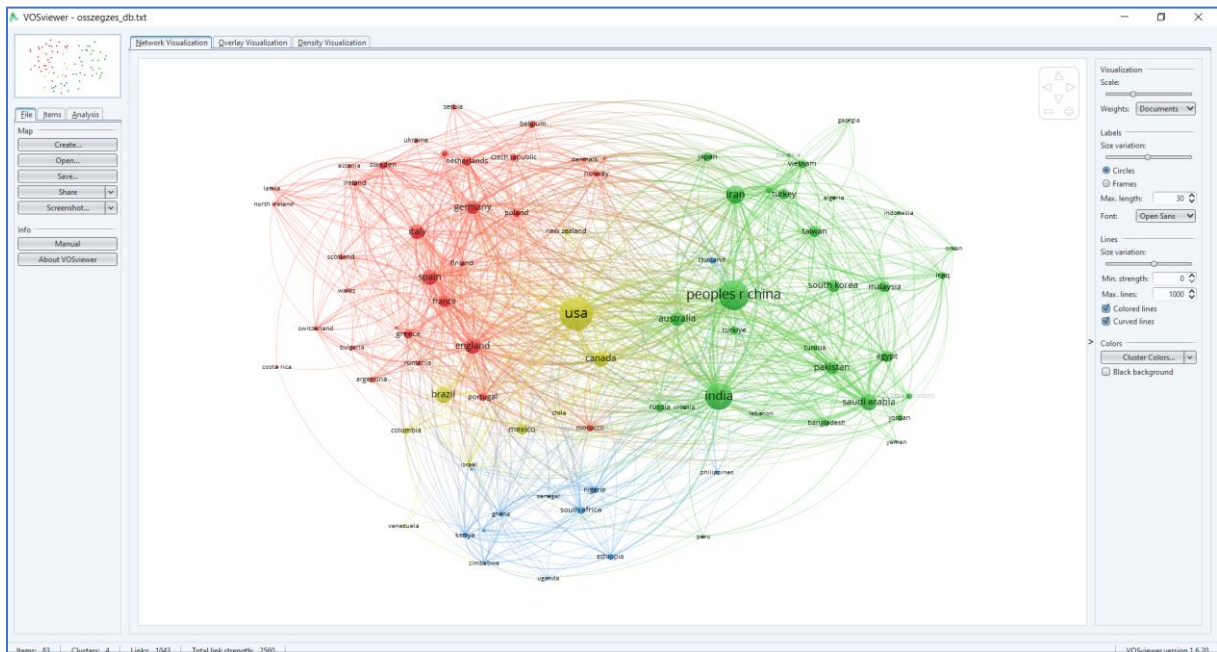


A co-citation (magyarul egyidejű hivatkozás) egy olyan fogalom a tudományos kutatásokban, amely arra utal, hogy két vagy több tudományos cikk vagy forrás egyszerre hivatkozik ugyanarra a harmadik forrásra. Ez azt jelenti, hogy az adott cikkek, kutatások vagy szerzők közötti kapcsolatot vagy kapcsolódási mintát észlelhetünk, mivel ugyanazokat a forrásokat használják fel az adott témában.

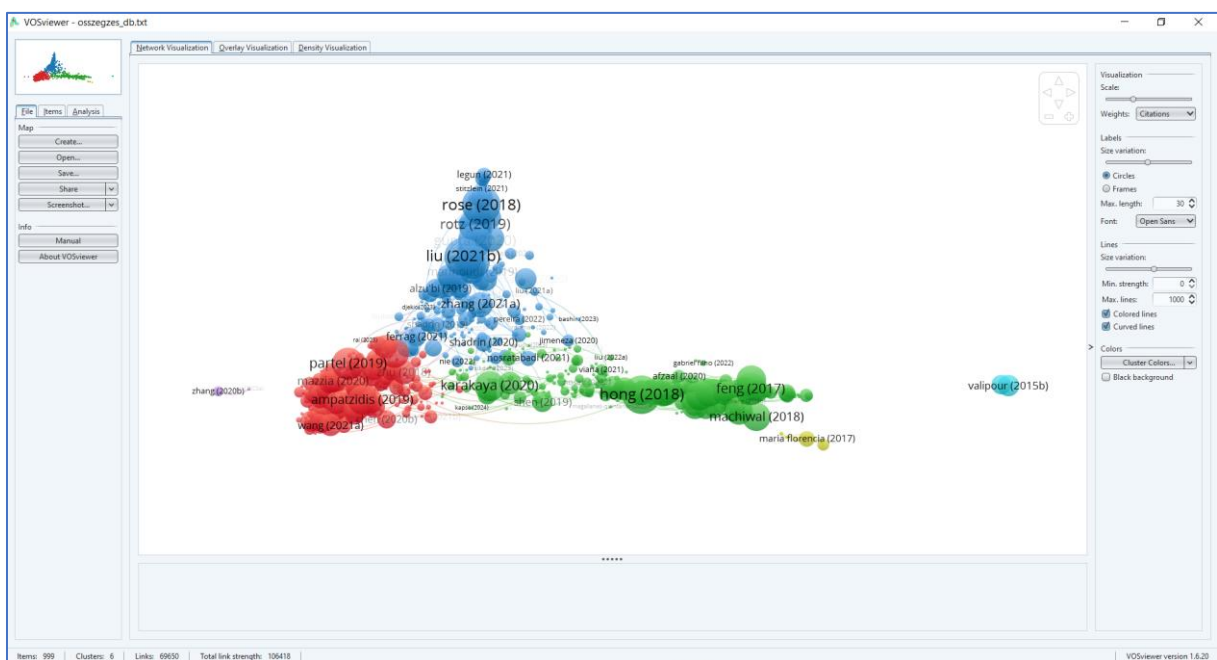
A VOSviewer-ben a co-citation elemzés lehetővé teszi a tudományos cikkek közötti egyidejű hivatkozások vizualizációját és elemzését. Az általában a VOSviewer hálózati térképén vagy grájfján jeleníthető meg, ahol a csomópontok a cikkeket vagy szerzőket, az élek pedig a közöttük lévő co-citation kapcsolatokat jelképezik.

A co-citation elemzés fontos eszköz lehet a tudományos kutatásokban, mivel segítségével azonosíthatók azok a cikkek vagy szerzők, amelyek között jelentős kapcsolat áll fenn egy adott témában. Ez lehetővé teszi az adott területen végzett kutatások és az egyes szerzők vagy cikkek közötti kapcsolatok felfedezését, és segíthet az új irányok vagy kutatási trendek azonosításában.

20. ábra: Az országok kapcsolata
(Forrás: saját munka VOSviewer-ben)



21. ábra: Hivatkozási háló
(Forrás: saját munka VOSviewer-ben)



Ugyanakkor erre a feladatra inkább a következőkben bemutatásra kerülő eszközt használjuk.

4.3.4 Adatelemzés CitNetExplorer-rel

22. ábra: CitNetExplorer logója
(Forrás: www.citnetexplorer.nl)



4.3.4.1 A CitNetExplorer szoftver jellemzői és alkalmazási lehetőségei a tudományos kutatásokban

A CitNetExplorer egy olyan szoftveres eszköz, amelyet kifejezetten a tudományos publikációk hálózati elemzésére és vizualizációjára terveztek. Ez a szoftver lehetővé teszi a tudományos kutatásokban használt hivatkozási hálózatok elemzését és megértését, és segít az adott tudományos területek közötti kapcsolatok és trendek feltárásában.

Funkciók és tulajdonságok

Hivatkozási hálózatok elemzése: A CitNetExplorer segítségével azonosíthatóak és elemezhetőek a tudományos publikációk közötti hivatkozási kapcsolatok. Ezáltal megérthetővé válnak az adott tudományterületek vagy témák közötti kapcsolatok, a fejlődési irányok.

Hálózati vizualizációk: A szoftver különböző grafikus formában lehetővé teszi a hivatkozási hálózatok vizualizációját, melyek segítenek a kapcsolódási mintázatok és trendek könnyebb feltárásában.

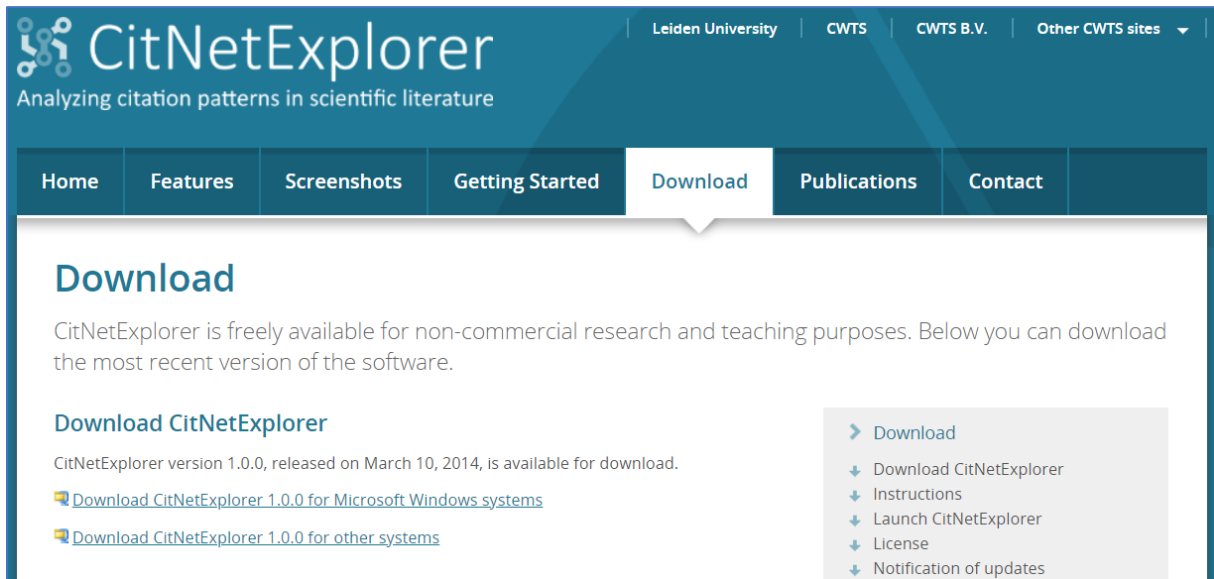
Csoportosítási és klaszterezési funkciók: A CitNetExplorer lehetővé teszi a hivatkozási hálózatok csoportosítását és klaszterezését különböző módszerek alapján. Ez segít az összetartozó elemek azonosítását, és azok klaszteres vizuális megjelenítését.

Alkalmazási lehetőségek a tudományos kutatásokban:

- Tudományos publikációk közötti kapcsolatok és trendek feltárása.
- Az adott tudományterületek vagy témák közötti kapcsolódási mintázatok azonosítása.
- Tudományos közösségek vagy kutatócsoportok közötti kapcsolatok vizsgálata.
- Az új irányok és lehetőségek azonosítása a tudományos kutatások terén.

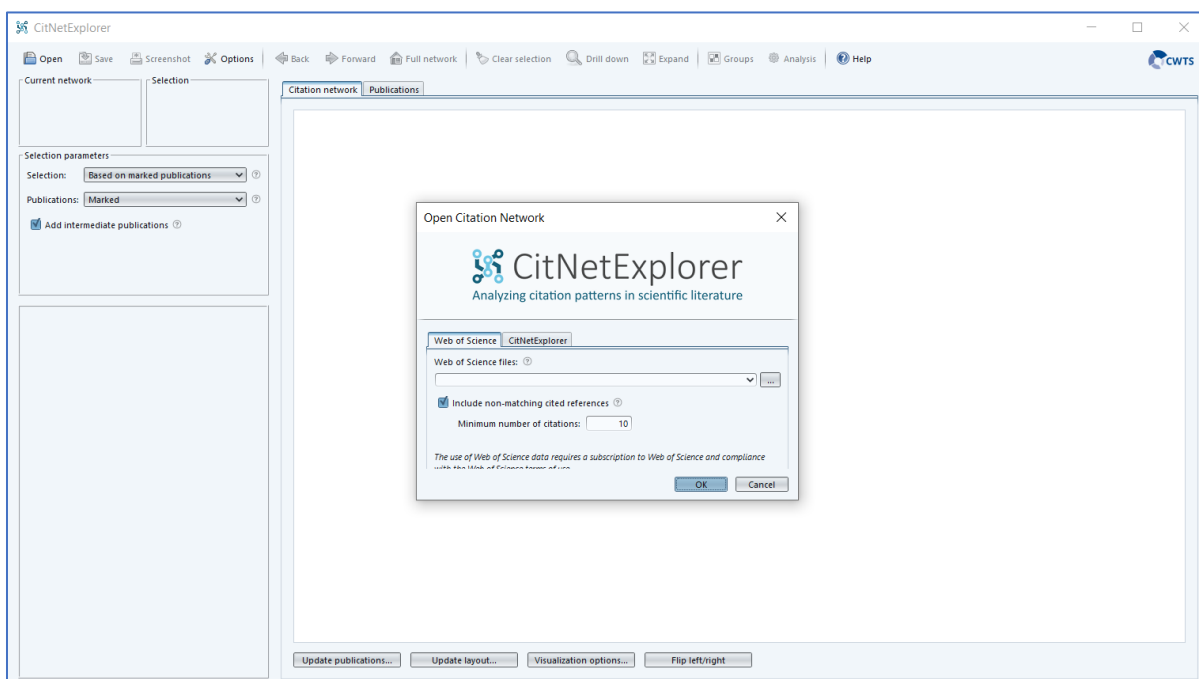
Némi megszorítás, hogy kizárólag a Web of Science-ből tud dolgozni.

23. ábra: A CitNetExplorer letöltési felülete
(Forrás: <https://www.citnetexplorer.nl/download>)



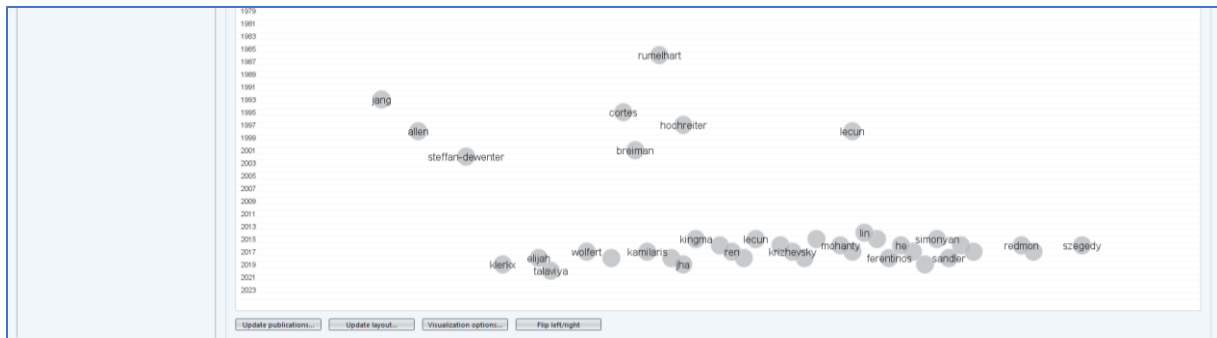
A program rendkívül érzékeny arra, ha nincsen megadva a publikáció éve. Viszont nagy adatbázisoknál előfordulhat, hogy nem minden esetben adják meg a publikáció évét. Akkor ez az eszköz nem fog működni! Ezért egyenként kell az előforduló évszámokból megadni a kívánt értékeket. Akár az összes is bepipálható, mellyel egyúttal kizárható az évszám nélküli rekord behúzása a vizsgálatba. A 24. ábra mutatja a szakirodalmi információs fájl feltöltési felületét.

24. ábra: A CitNetExplorer adatfájl bekérő felülete
(Forrás: saját munka CitNetExplorer-ben)



CitNetExplorer kiválóan alkalmas arra, hogy ontológia, episztemológia kérdésekre keressünk válaszokat: meghatározható a tudásteremtés eredője, azonosítható az a szerző, az a cikk, ami annak a tématerületnek meghatározó kiindulópontjaként működik a későbbi publikációk hivatkozásai alapján. (25. ábra)

25. ábra: Hivatkozási kapcsolatok vizualizációja a különböző szerzők között
(Forrás: saját munka CitNetExplorer-ben)

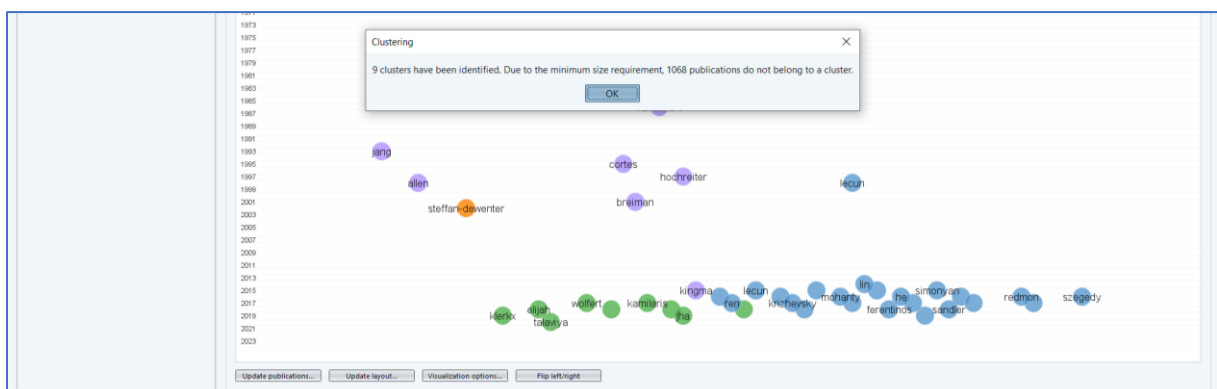


Az alkalmazandó menükombináció: Analysis/Clustering: Merge small clusters

A tudományos kutatásokban gyakran találkozunk olyan helyzetekkel, amikor fontos az adott kutatási területek vagy témák közötti kapcsolatok és klaszterek azonosítása. A CitNetExplorer klaszterezés funkciója lehetővé teszi a tudományos publikációk közötti kapcsolódási mintázatok és klaszterek azonosítását, ami segíti a kutatókat az adott tudományos területek vagy témák megértésében és vizualizációjában.

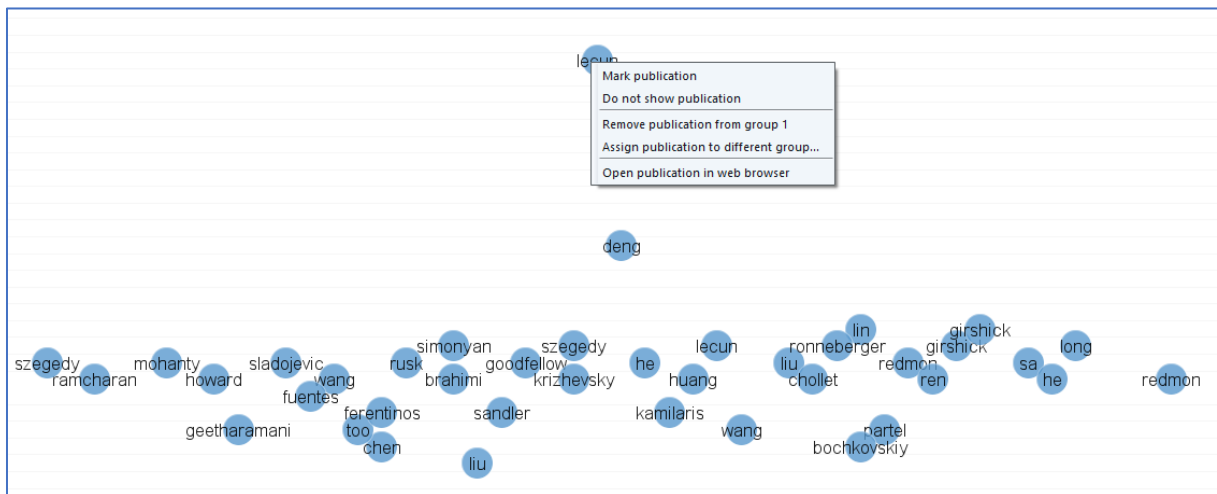
A felhasznált hivatkozási struktúra alapján a CitNetExplorer 9 klaszterre bontotta az adathalmazt (26. ábra). Ezt követően az egyes klaszterek külön-külön is vizsgálhatóak.

26. ábra: A CitNetExplorerben azonosított klaszterek megjelenése
(Forrás: saját munka CitNetExplorer-ben)



Így minden klaszterben azonosítható a tématerületet megalapozó cikk, melyet egyetlen kattintással elérhetünk, ha már rendelkezik elektronikus dokumentumazonosítóval. (27. ábra és 28. ábra)

27. ábra: A meghatározó irodalom elektronikus elérésének módja
(Forrás: saját munka CitNetExplorer-ben)



28. ábra: Az alap forrásként azonosított cikk elektronikus megjelenése
Forrás: <https://ieeexplore.ieee.org/document/726791>

Journals & Magazines > Proceedings of the IEEE > Volume: 86 Issue: 11

Gradient-based learning applied to document recognition

Publisher: IEEE [Cite This](#) [PDF](#)

Y. Lecun ; L. Bottou ; Y. Bengio ; P. Haffner [All Authors](#)

32117 Cites in Papers	212 Cites in Patents	97838 Full Text Views
------------------------------------	-----------------------------------	------------------------------------

[R](#) [Share](#) [CC](#) [Folder](#) [Bell](#)

Abstract

Abstract: Multilayer neural networks trained with the back-propagation algorithm constitute the best example of a successful gradient based learning technique. Given an appropriate network architecture, gradient-based learning algorithms can be used to synthesize a complex decision surface that can classify high-dimensional patterns, such as handwritten characters, with minimal preprocessing. This paper reviews various methods applied to handwritten character recognition and compares them on a standard handwritten digit recognition task. Convolutional neural networks, which are specifically designed to deal with the variability of 2D shapes, are shown to outperform all other techniques. Real-life document recognition systems are composed of multiple modules including field extraction, segmentation recognition, and language modeling. A new learning

[Authors](#)
[References](#)
[Citations](#)
[Keywords](#)
[Metrics](#)

4.3.5 Adatelemzés Bibliometrix-szel

29. ábra: A Bibliometrix logója
(Forrás: www.bibliometrix.org)



Bibliometrix egy olyan R nyelvhez fejlesztett csomag, amelyet kutatók és bibliográfiai adatelemzők használnak a tudományos publikációk és azok hivatkozásainak bibliometriai

elemzésére. A Bibliometrix segítségével lehetőség van az olyan metrikák kiszámítására és vizualizálására, mint például a publikációk időbeli eloszlása, a szerzők közötti együttműködések hálózata, a hivatkozások általános mintázata és a tudományterületek közötti kapcsolatok.

A Bibliometrix az R nyelvhez fejlesztett csomag, ami azt jelenti, hogy az R programozási nyelvet használja a működéséhez, ott futtathatóak a lefejlesztett parancsai. Az R egy nagyon népszerű statisztikai eszköz a tudományos közösségekben. Ennek eredményeként a Bibliometrix használata széles körben elterjedt a tudományos publikációk és kutatások bibliometriai elemzésére.

A Bibliometrix számos funkciót kínál az elemzéshez és a vizualizációhoz, beleértve a következőket:

- Publikációs adatok importálása és előkészítése.
- Publikációk hálózatának elemzése (pl. szerzői hálózatok, hivatkozási hálózatok stb.).
- Publikációk időbeli elemzése és trendek vizsgálata.
- Tudományterületek közötti kapcsolatok elemzése.
- Publikációk jelentőségének és befolyásának mérésére szolgáló mutatók számítása (pl. impact factor, h-index stb.).

Összességében a Bibliometrix egy hatékony eszköz a tudományos publikációk és kutatások bibliometriai elemzésére és vizualizációjára az R programozási nyelv környezetében.

4.3.5.1 RStudio: biblioshiny

30. ábra: Az RStudio logója
(Forrás: www.rstudio.com)



Némi kompromisszum árán az R programmal könnyen dolgozhatunk az RStudio szoftver segítségével. (30. ábra)

Néhány ezer adatig jól használható a Bibliometrix web-es felülete, melyet a 'biblioshiny()' paranccsal tudunk elindítani (31. ábra). Ebben az esetben nem a felhasználó gépén, hanem

egy szerveren futnak le a számítások, ami azonban az adatkapcsolat függvényében lassulást jelenthet.

31. ábra: A 'biblioshiny' parancs kiadása
(Forrás: saját munka RStudio-ban)



```
R 4.3.2 · ~/
Help us to keep Bibliometrix and Biblioshiny free to download and use by contributing with a small
donation to support our research team (https://bibliometrix.org/donate.html)

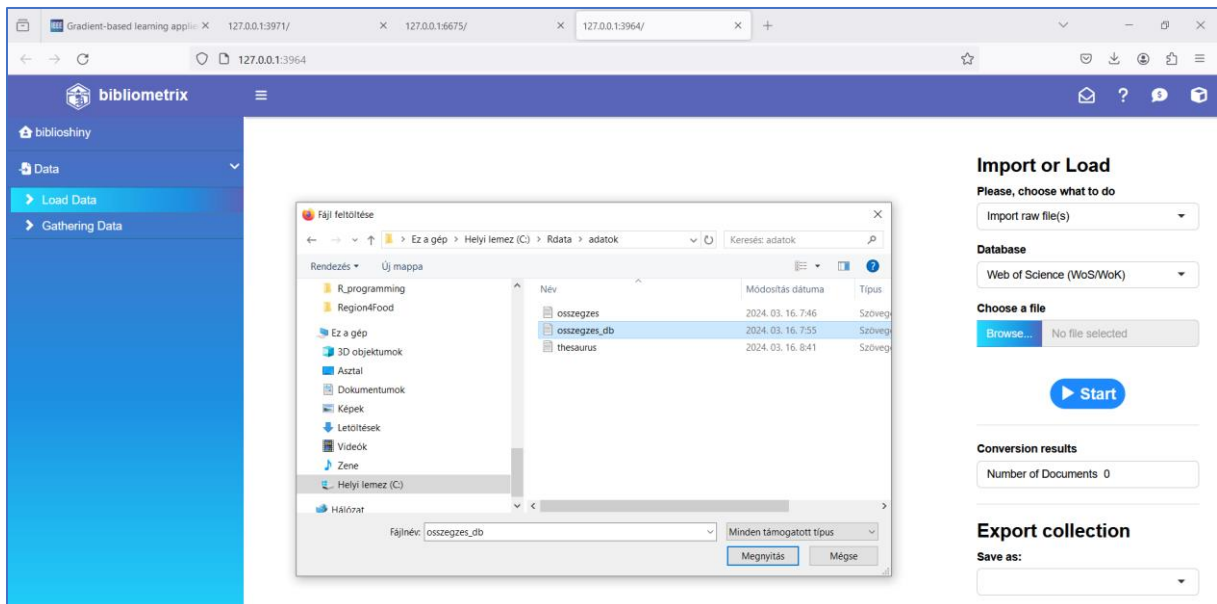
To start with the Biblioshiny app, please digit:
biblioshiny()

Warning message:
package 'bibliometrix' was built under R version 4.3.3
> biblioshiny()
```

A parancs kiadását követően elindul a szerverrel való kapcsolatfelvétel. (32. ábra)

A "biblioshiny" egy olyan R nyelvhez fejlesztett csomag, amely egy interaktív Shiny alkalmazást biztosít a Bibliometrix csomag felhasználói számára. A Shiny szintén egy R nyelvhez készült csomag, amely lehetővé teszi interaktív webes alkalmazások létrehozását az R kód segítségével.

32. ábra: A szakirodalmi információs fájl beolvasása Bibliometrix-be
(Forrás: saját munka a bibliometrix-ben)



A biblioshiny használatával a felhasználók könnyen létrehozhatnak interaktív grafikonokat, táblázatokat és egyéb vizualizációkat a Bibliometrix csomaggal végzett elemzéseik eredményeinek megjelenítésére. Ez lehetővé teszi a kutatók számára, hogy könnyen

megosszák az eredményeket másokkal, és interaktív módon vizsgálják meg a publikációk bibliometriai jellemzőit és trendjeit.

A biblioshiny segítségével a felhasználók egy felhasználóbarát felhasználói felületen keresztül hajthatják végre a Bibliometrix elemzéseket, ami lehetővé teszi az elemzések gyorsabb és hatékonyabb elvégzését. A beolvasással itt is egy dataframe készül, mely 86 oszlopból és több százezer sorból áll. A folyamat végén megjelenő ablak szemléletesen ismerteti a beolvasott adatok minőségét, mely jelen esetben egészen kiváló jellemzőket mutat. (33. ábra)

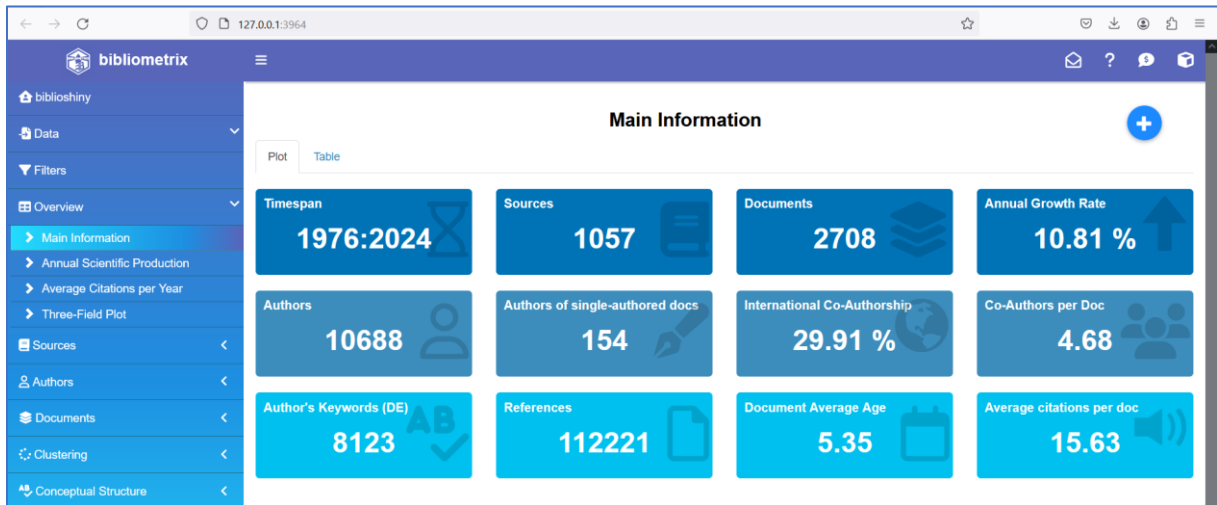
33. ábra: A Bibliometrixbe beolvasott adatok minősége
(Forrás: saját munka a bibliometrix-ben)

Completeness of bibliographic metadata - 2708 documents from Isi				
Metadata	Description	Missing Counts	Missing %	Status
AU	Author	0	0.00	Excellent
DT	Document Type	0	0.00	Excellent
SO	Journal	0	0.00	Excellent
LA	Language	0	0.00	Excellent
PY	Publication Year	0	0.00	Excellent
WC	Science Categories	0	0.00	Excellent
TI	Title	0	0.00	Excellent
TC	Total Citation	0	0.00	Excellent
AB	Abstract	5	0.18	Good
RP	Corresponding Author	24	0.89	Good
CR	Cited References	28	1.03	Good
C1	Affiliation	52	1.92	Good
DI	DOI	170	6.28	Good
DE	Keywords	262	9.68	Good
ID	Keywords Plus	574	21.20	Poor

A felület nagyon jó menü- és help-rendszerrel működik, így az egyéni megismerésre, feltérképezésre igen jó lehetőség mutatkozik. A legfontosabb információk az #Overview\Main Information menüpont alatt érhetőek el. (34. ábra)

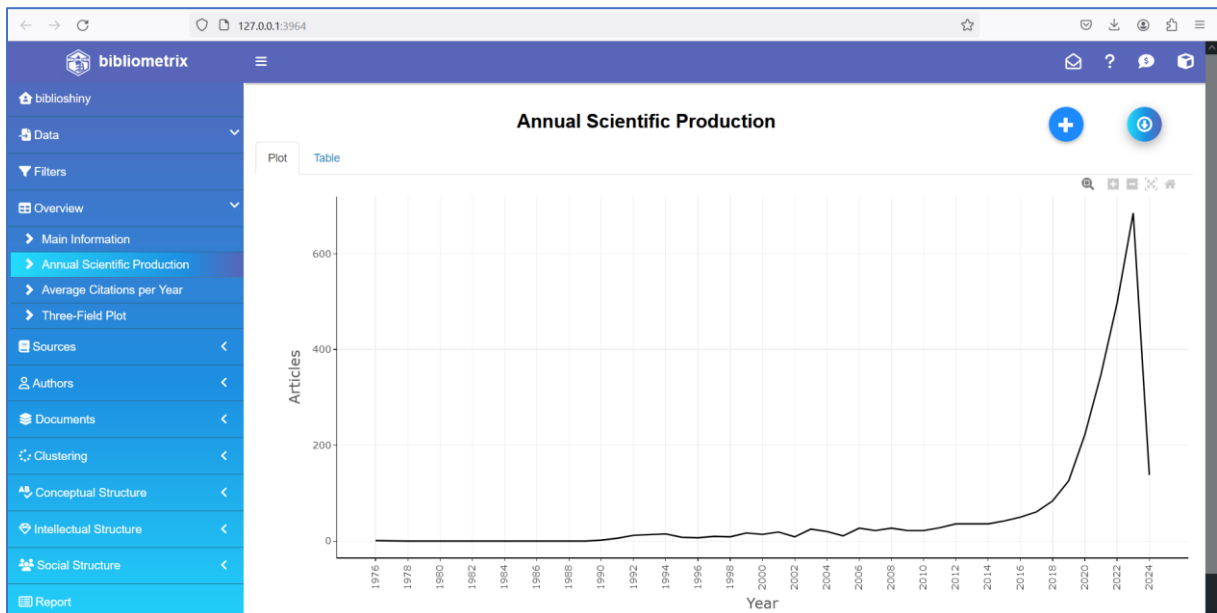
A kapott ábra könnyen értelmezhető: a cikkek előfordulása 1976-2024 közöttre tehető. 1057 forrásból összesen 2708 dokumentum adatait dolgozta fel a rendszer. És így tovább értelmezhetőek az egyes „csempék”.

34. ábra: A legfontosabb adatok kinyerése bibliometrix-ben
(Forrás: saját munka a bibliometrix-ben)



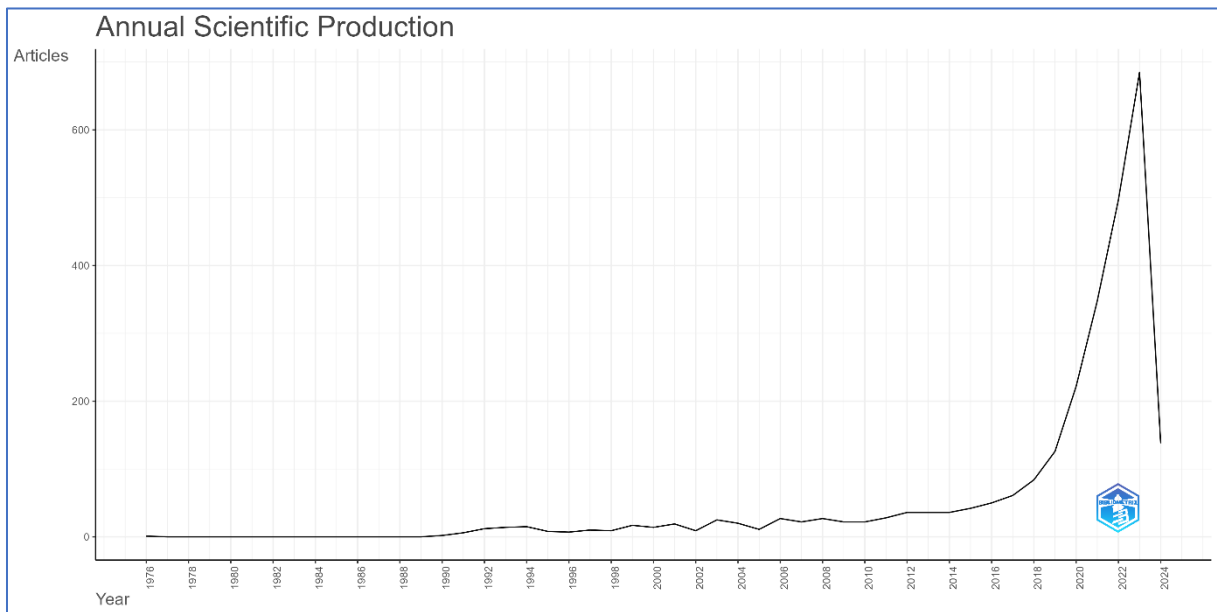
Ugyanazokat az információkat kapjuk, mintha R-ben a 'summary' parancsot adnánk ki, csak kicsit kényelmesebb módon, illetve a kapott eredményeket is könnyebben olvassa a szem. Nagyon gyorsan és könnyen elérhetőek a leggyakrabban használt vizualizációs technikák, például az évenkénti publikációk száma. (35. ábra)

35. ábra: Az évenkénti publikációk mennyisége a bibliometrix-ben
(Forrás: saját munka a bibliometrix-ben)



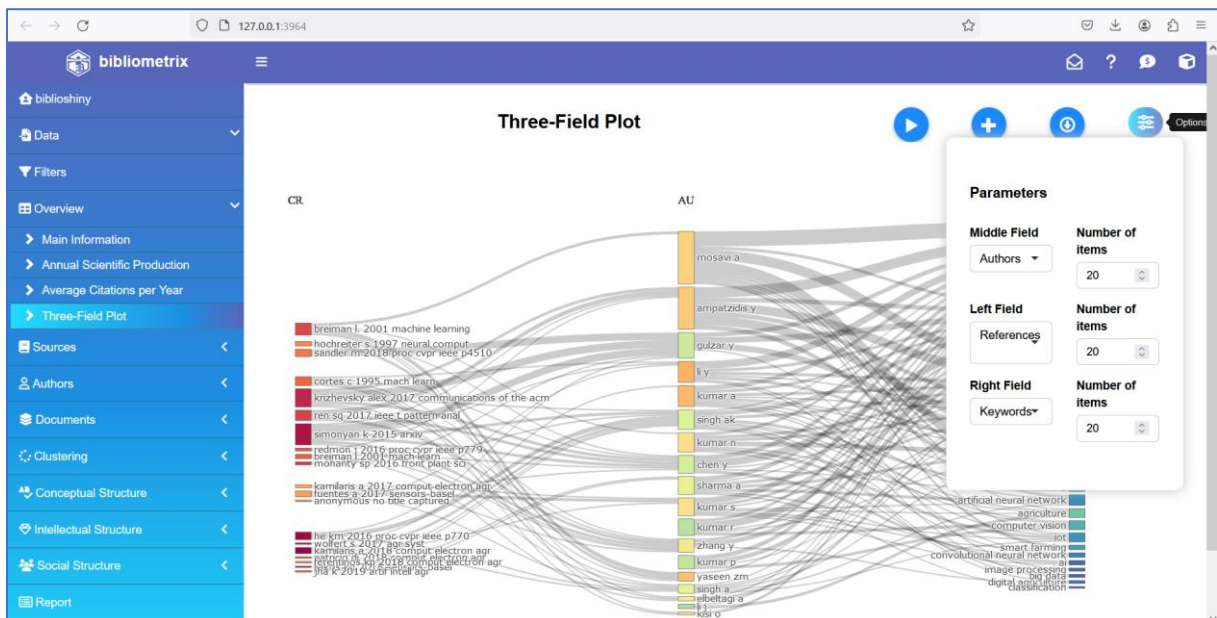
A felület erősen támogatja a tudományos írásművek készítéséhez szükséges ábrák exportját is. (36. ábra)

36. ábra: Ábra export a bibliometrix-ből
(Forrás: saját munka a bibliometrix-ben)



Rendkívül népszerű a mai szakirodalomban a háromdimenziós bibliográfiai összefüggések bemutatására használt Three-Field Plot diagram. A Bibliometrix-ben ez könnyen elérhető, egyszerűen paramétrezhető, konferencia előadó, szakértő keresésekor is jól használható. (37. ábra)

37. ábra: Háromdimenziós bibliográfiai összefüggések a Bibliometrix-ben
(Forrás: saját munka a bibliometrix-ben)



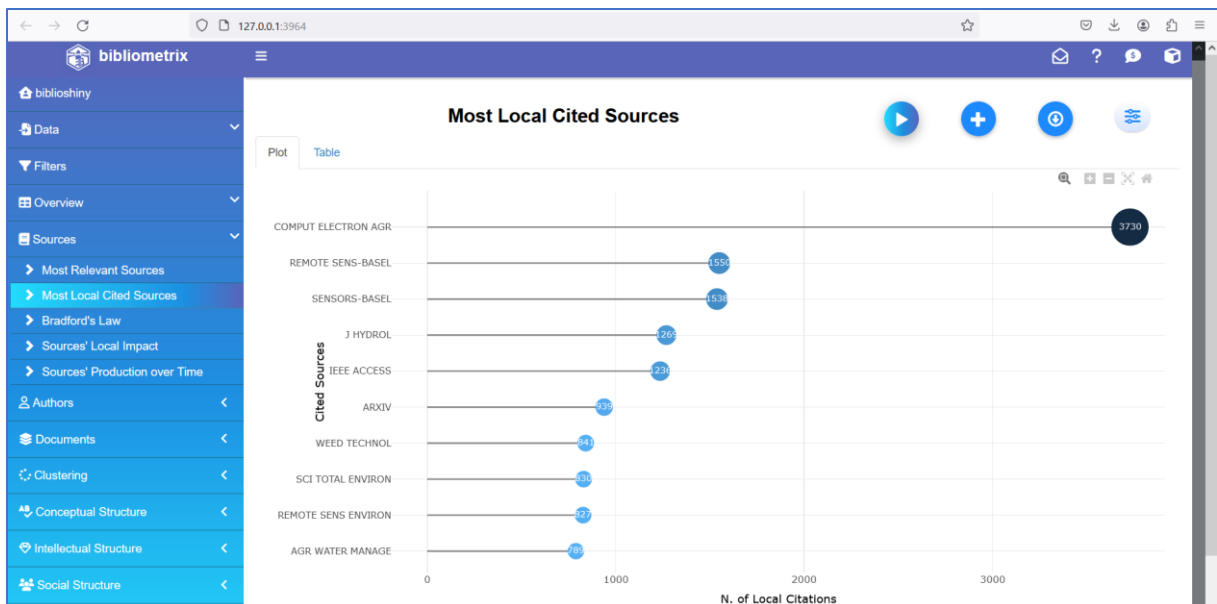
Rendkívül könnyen meghatározhatóak azok a folyóiratok, amelyekben a legtöbb, a tématerületet érintő cikk jelenik meg (38. ábra). Azonban ez nem azonos a legtöbbet

hivatkozott forrásokkal. Jelen esetben is jól látszik, hogy milyen markáns eltérés tapasztalható ezen a téren (39. ábra).

38. ábra: A legtöbb cikket felmutató források
(Forrás: saját munka a bibliometrix-ben)



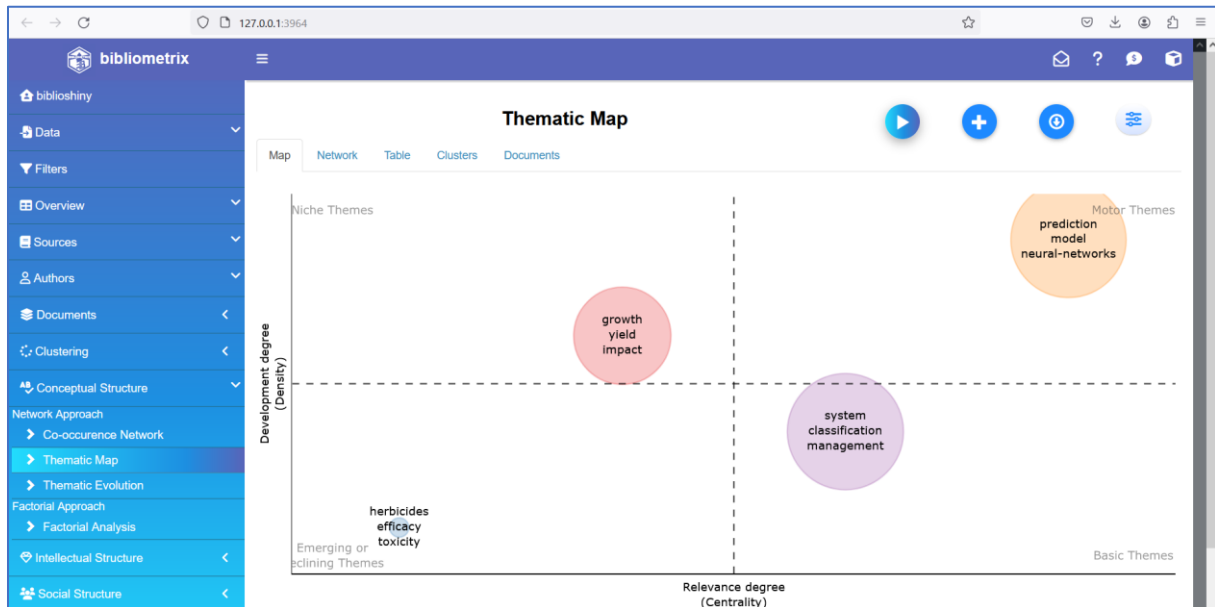
39. ábra: A legtöbb hivatkozást felmutató források
(Forrás: saját munka a bibliometrix-ben)



#Conceptual Structure \ Thematic Map menüvel egy olyan klaszterezést érhetünk el, amely három fő ismérv mentén vizsgálja a cikkek összetartozását: kiket idéznek, kik hivatkoznak rá, milyen kulcsszavak jelennek meg benne. Ez alapján egy multidimenziós tér jön létre, melyet

kétdimenziós megjelenítési formába transzformál, amihez jelentéstartalmat is rendel.
(40. ábra)

40. ábra: Tematikus térkép a Bibliometrixben
(Forrás: saját munka a bibliometrix-ben)



Motor, centrális témák sok ide mutató hivatkozással (centrality), sok egymás közötti hivatkozással (density). (Motor Themes)

Specifikus, niche témák: kevesen hivatkoznak rájuk, inkább csak saját köreiken belül történik hivatkozás, intenzív párbeszéd. (Niche Themes)

Alacsony érdeklődés mind kívülről, mind belülről a szakmai közösséget nézve. Ennek két oka lehet: vagy már lecsengett az adott tématerület aktualitása, vagy most van felszálló pályán. (Emerging or declining Themes)

Alapvető témák jellemzője, hogy kívülről sok hivatkozást kapnak, de az egymás közötti párbeszéd intenzitása alacsony. (Basic Themes)

Különböző témák közötti kapcsolat bemutatására faktor analízist is készíthetünk a Bibliometrix-ben, melyet dendogramos megjelenítés szemléltet (41. ábra)

5 Eredmények

Diplomamunkám során áttekintettem azokat a bibliográfiai adatvizualizációs eszközöket, szoftvereket, melyekkel leginkább lehet találkozni a legfrissebb tudományos munkákban. Ehhez a Magyar Agrár- és Élettudományi Egyetem Gazdaság- és Regionális Tudományok Doktori Iskola valamint a Debreceni Egyetem Ihrig Károly Gazdálkodás- és Szervezéstudományok Doktori Iskola disszertációi adtak alapot.

Részletesen az alábbi három alkalmazást vizsgáltam:

- VOSviewer
- CitNetExplorer
- bibliometrix

A diplomamunkámat alkalmazott informatikai dolgozatnak tekintem. Ennélfogva az eredmények tárgyalása több általánosabb megállapítást tartalmaz, mint egy szoftver- vagy alkalmazásfejlesztésen alapuló munka. Ugyanakkor kijelenthető, hogy az alkalmazott informatikai kutatások az egyik legfontosabb területei a modern tudományos és technológiai fejlődésnek. Ezen a területen végzett kutatások és fejlesztések jelentős és közvetlen hatással vannak mindennapi életünkre, a gazdasági javak előállítására és a társadalomra. Számos okból fontos és értékes az alkalmazott informatikai kutatás:

1. **Technológiai innováció elősegítése:** Az alkalmazott informatikai kutatás olyan új technológiák és megoldások fejlesztésére összpontosít, amelyek jelentősen hozzájárulnak a technológiai innovációhoz. Ezek a fejlesztések lehetővé teszik az új alkalmazások létrehozását, amelyek javítják az emberek életminőségét és hatékonyabban működő vállalkozásokat és szervezeteket eredményeznek.
2. **Társadalmi kihívásokra adott válaszok:** Az alkalmazott informatikai kutatás segít megoldani a társadalmi kihívásokat és problémákat. Ilyenek lehetnek például az egészségügyi ellátás javítása, a környezetvédelem, a városfejlesztés vagy a társadalmi egyenlőtlenségek csökkentése. Az új technológiai megoldások és alkalmazások segítenek hatékonyabbá tenni az ilyen területeken folytatott munkát.
3. **Gazdasági fejlődés elősegítése:** Az informatikai kutatás és fejlesztés kulcsfontosságú a gazdasági fejlődés előmozdításában. Az új technológiák és termékek létrehozása új munkahelyeket teremt, új vállalkozásokat hoz létre és hozzájárul a gazdasági növekedéshez és versenyképességhez.

4. **Jövőbeli kihívásokra való felkészülés:** Az alkalmazott informatikai kutatás segít felkészülni és válaszolni a jövőbeni kihívásokra és lehetőségekre, például az automatizáció, mesterséges intelligencia, adatvédelem vagy a kiberbiztonság terén. Az ilyen kutatások elősegítik a társadalom és az ipar rugalmasságát és alkalmazkodóképességét a változó környezethez.
5. **Tudományos előrelépések:** Az alkalmazott informatikai kutatás segíti a tudományos és technológiai előrelépéseket, új ismeretek és módszerek létrehozásával. Ezek az előrelépések hozzájárulnak a tudományág fejlődéséhez és a tudományos közösség általános tudásának bővítéséhez.

A diplomadolgozatomban ez utóbbi területhez, a bibliográfiai adatok elemzéséhez használtam különböző szoftvereket. Ezek előnyének és hasznosságának bemutatása fontos fókuszterületet képviseltek a dolgozatban, de az alkalmazhatóság korlátjaira is kitértem.

Az ilyen szoftverek használata az adatelemzés hatékonyságát és pontosságát is növelik, és számos módon hozzájárulnak a kutatási folyamatokhoz:

Először is, ezek a szoftverek lehetővé teszik a nagy mennyiségű adatok gyors és hatékony feldolgozását. A manuális adatelemzés hosszadalmas és időigényes lehet, míg ezek a szoftverek automatizált funkciói lehetővé teszik a nagy adatmennyiségek gyors feldolgozását és elemzését, így időt és erőforrásokat takarítanak meg.

Másodszor, ezek a szoftverek többféle elemzési módszert és technikát kínálnak, amelyek segítenek a kutatóknak a változatos kutatási kérdések megválaszolásában. Például lehetőség van tematikus elemzésre, tartalomelemzésre, hálózatelemzésre és más analitikai módszerekre.

Harmadszor, ezek a szoftverek általában rugalmasak és testreszabhatóak a kutatók igényei szerint. Lehetőség van a különböző adatforrások importálására és integrálására, valamint a különböző paraméterek és beállítások testreszabására az adatelemzés céljainak megfelelően. Végül, ezek a szoftverek általában könnyen kezelhetők és felhasználóbarát felülettel rendelkeznek, amely lehetővé teszi még azok számára is a használatukat, akik kevésbé jártasak az adatelemzésben.

Összességében tehát az elemzéshez használt szoftverek jelentős előnyöket kínálnak a kutatók számára, és fontos szerepet játszanak a kutatási folyamat hatékonyságának és eredményességének növelésében.

6 Következtetések és javaslatok

Kutatásaim során tapasztaltak alapján megfogalmazható néhány lehetőség és iránymutatás, amelyek segíthetnek a bibliográfiai elemző szoftverek hatékonyabb és eredményesebb alkalmazásában. Ezek a javaslatok az alábbiakban foglalhatók össze:

1. **Forrásadatok optimalizálása:** Fontos, hogy a kutatók olyan forrásadatokat válasszanak és használjanak, amelyek megfelelnek a kutatási céloknak és igényeknek. Az optimális forrásadatok kiválasztása teszi lehetővé a releváns és megbízható elemzést.
2. **Adatfeldolgozás módszereinek alkalmazása:** A felkészülten megválasztott adatfeldolgozási módszerek segíthetnek abban, hogy a kutatók hatékonyan kezeljék és elemezzék a nagy mennyiségű bibliográfiai adatokat. Továbbá a megfelelő feldolgozási módszerek kiválasztása lehetővé teszi az adatelemzés hatékonyságának és pontosságának növelését is.
3. **Összehasonlító elemzés:** Fontos, hogy a kutatók alkalmazzák az összehasonlító elemzés módszereit a különböző szoftverek között, azaz legyenek tisztában az egyes alkalmazások előnyeivel és korlátaival. Az összehasonlító vizsgálatok segíthetnek abban, hogy a kutatók objektív és megbízható véleménnyel rendelkezzenek az alkalmazott szoftverek teljesítményéről.
4. **Felhasználói visszajelzések gyűjtése:** Fontos, hogy a fejlesztők gyűjtsék és értékeljék a felhasználói visszajelzéseket a kínált szoftverekről. A felhasználói visszajelzések segíthetnek azonosítani azokat a területeket, ahol javításokra vagy fejlesztésekre van szükség, és segíthetnek abban, hogy az alkalmazott szoftverek még hatékonyabbak legyenek a jövőben. Ezen általánosnak tűnő mondatok mögött sok olyan megtapasztalás áll, amikor egészen apró ismereteken múlik egy vizsgálat sikere vagy totális félrecsúszása.
5. **Képzés és támogatás:** Fontos, hogy a kutatók részesüljenek megfelelő képzésben és támogatásban az alkalmazott szoftverek használatához. A megfelelő képzés és támogatás lehetővé teszi, hogy a kutatók a lehető leghatékonyabban kihasználják az alkalmazott szoftverek által nyújtott lehetőségeket és funkciókat.

Ezek a javaslatok és iránymutatások segíthetnek abban, hogy a bibliográfiai elemző szoftverek hatékonyabban és eredményesebben legyenek alkalmazhatók a kutatási folyamatokban.

7 Összefoglalás

A diplomadolgozatom célja az volt, hogy feltárja és bemutassa a bibliográfiai elemző szoftverek jelentőségét és alkalmazhatóságát a tudományos kutatásokban. Az elemzés során számos fontos témát érintettem, amelyek alapján megállapítható, hogy ezek a szoftverek kulcsfontosságú szerepet játszanak a kutatási folyamatok hatékonyabbá és eredményesebbé tételében.

A dolgozat készítése során megvizsgáltam a bibliográfiai elemző szoftverek funkcionalitását és hatékonyságát, bemutattam azok előnyeit és korlátait. A kutatás eredményeként megállapítottam, hogy ezek a szoftverek képesek jelentős mértékben segíteni a kutatókat az adatok gyors és hatékony feldolgozásában, valamint a releváns szakirodalom feltárásában.

Emellett rámutattam az alkalmazott informatikai módszerek és technológiák jelentőségére a bibliográfiai adatok elemzésében, és kritikusan megvizsgáltam azokat az előnyöket és korlátokat, amelyekkel ezek a módszerek és eszközök járnak a tudományos kutatásban.

A kutatási részben kifejtettem a Web Of Science tudományos bibliográfiai portál kereső funkciójának részleteit, a keresési algoritmus sajátosságait. Bemutattam a szakirodalmi információk fájl létrehozásának lépéseit, annak fontosságát és buktatóit. Ezt követően több bibliográfia elemző szoftver alkalmazását mutattam be összehasonlító elemzés, valamint egy komplex eszközszer összeállítása céljából.

A dolgozat alapján arra a következtetésre jutottam, hogy a bibliográfiai elemző szoftverek nagy potenciált hordoznak a kutatási folyamatok hatékonyabbá és eredményesebbé tételében, azonban fontos felismerni és megérteni az ezekkel a szoftverekkel kapcsolatos kihívásokat, korlátokat is; és folyamatosan fejleszteni kell a velük kapcsolatos tudásunkat.

Végső soron remélem, hogy a diplomadolgozat eredményei hozzájárulnak majd a kutatók és szakemberek számára az információelemzés terén folytatott tevékenységük hatékonyabb és eredményesebb végrehajtásához, és elősegíti az Egyetemünkön, doktori iskoláinkban folyó kutatói munkát!

8 Irodalomjegyzék

- Aria, M. & Cuccurullo, C., 2017. bibliometrix: An R-tool for comprehensive science mapping. *Journal of Informetrics*, 2017(11), pp. 959-975.
- Batagelj, V. & Cerinsek, M., 2013. On bibliographic networks. *Scientometrics*, 2013(96), pp. 845-864.
- Chen, C., 2017. Science Mapping: A Systematic Review of the Literature. *Journal of Data and Information Science*, 2(2), pp. 1-40.
- Chen, H., Tsang, Y. & Wu, C., 2023. When text mining meets science mapping in the bibliometric analysis: A review and future opportunities. *International Journal of Engineering Business Management*, 15. kötet, pp. 1-15.
- Cobo, M., Lopez-Herrera, A., Herrera-Viedam, E. & Herrera, F., 2011. Science Mapping Software Tools: Review, Analysis, and Cooperative Study Among Tools. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 62(7), pp. 1382-1402.
- Cobo, M., Lopez-Herrera, A., Herrera-Viedma, E. & Herrera, F., 2010. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, 2011(5), pp. 146-166.
- Giustin, D. & Boulus, M. N., 2013. Google Scholar is not enough to be used alone for systematic reviews. *Online J Public Health Inform*, 5(2), p. 214.
- Glanzel, W., Moed, H. F., Schmoch, U. & Thelwall, M., 2019. *Handbook of Science and Technology Indicators*. Svájc: Springer.
- Kamada, T. & Kawai, S., 1989. An Algorithm for Drawing General Undirected Graphs. *Information Processing Letters*, 31(1), pp. 7-15.
- Qu, H., Nordin, N. A., Tsong, T. B. & Feng, X., 2023. A Bibliometrics and Visual Analysis of Global Publications for Cognitive Map. *IEEE Access*, 11. kötet, pp. 52824-52839.
- Smiraglia, R. P. & Leazer, G. H., 1999. Derivative Bibliographic Relationships: The Work Relationship in a Global Bibliographic Database. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE*, 50(6), pp. 493-504.
- Takács, M., 2014. *Fuzzy kognitív térképek*. Szabadka, Újvidéki Egyetem, pp. 107-117.
- van Raan, A. F., 2017. Patent Citations Analysis and Its Value in Research Evaluation: A Review and a New Approach to Map Technology-relevant Research. *Journal of Data and Information Science*, 2(1), pp. 13-50.
- Vine, R., 2006. Google Scholar. *J Med Libr Assoc.*, 2006(94), pp. 97-99.

Felhasznált szoftverek és adatbázisok:

bibliometrix www.bibliometrix.org)

CitNetExplore www.citnetexplorer.nl

Google Scholar (csak említésre került) www.scholar.google.com

MS Excel www.microsoft.com

Notepad++ <https://notepad-plus-plus.org>

Pajek (csak említésre került) <https://pajek.software.informer.com>

RStudio www.rstudio.com

Visual Studio Code <https://code.visualstudio.com/>

VOSviewer <https://www.vosviewer.com>

Web of Science: www.webofknowledge.com

9 Köszönetnyilvánítás

Kedves Barátaim!

- Bujdosó Zoltán
- Horváthné Kovács Bernadett
- Labátné Fercsik Katalin
- Lágymányosi Attila
- Lakner Zoltán
- Orova Lászlóné
- Sójáné Dux Ágnes
- Záborszki Norbert

Az Adattechnológus-adatelemző szakmérnök szakirányú továbbképzési szakot nem tudtam volna a segítségetek nélkül elvégezni.

Köszönöm szépen az iránymutatásokat, a türelmeteket és mindent, amivel támogattátok fejlődésemet, tanulmányaimat!

10 Mellékletek

'1. számú melléklet: Web of Science: All Export Field Tags

1.	A2	Other Abstract
2.	AA	Additional Authors
3.	AB	Abstract / BHTD Critical Abstract
4.	AD	Application Details and Date
5.	AE	Patent Assignee
6.	AF	Author Full Names
7.	AK	Abstract (Korean)
8.	AN	Accession Number or PubMedID
9.	AR	Article Number
10.	AU	Authors or Inventors
11.	AW	Item URL
12.	BA	Book Authors
13.	BD	Broad Descriptors or Broad Terms
14.	BE	Book Editor
15.	BF	Book Author Full Names
16.	BN	ISBN
17.	BP	Start Page
18.	BS	Book Series Subtitle
19.	C1	Addresses
20.	C2	Address (non-English)
21.	C3	Author Affiliations
22.	CA	Group Authors
23.	CC	Concept Codes or CABI Codes
24.	CE	Edition
25.	CH	Chemicals & Biochemicals
26.	CI	Derwent Compound Number
27.	CL	Conference Location
28.	CN	CAS Registry Numbers; Commercial Names; Chemical

29.	CO	CODEN
30.	CP	Cited Patent(s)
31.	CR	Cited References
32.	CT	Conference Title
33.	CY	Conference Date
34.	D2	Book DOI
35.	DA	Date of Export
36.	DC	Derwent Class Code(s)
37.	DE	Author Keywords; Descriptors
38.	DF	Date Filed or Submitted
39.	DI	DOI
40.	DL	DOI Link
41.	DM	Demography
42.	DN	DCR Number
43.	DP	Discipline; Diseases
44.	DS	Designated States
45.	DT	Document Type
46.	DY	Data Type
47.	EA	Early access date; Equivalent Abstract, Editor Address
48.	EC	Category
49.	ED	Editors
50.	EF	End of File
51.	EI	eISSN
52.	EM	E-mail Address
53.	EP	End Page
54.	ER	End of Record
55.	EY	Early access year
56.	FD	Further Application Details
57.	FN	File Name
58.	FP	Funding Name Preferred
59.	FS	Field of Search
60.	FT	Foreign Title

61.	FU	Funding Orgs
62.	FX	Funding Text
63.	GA	IDS Number
64.	GE	Geographic Data
65.	GI	Grant Information
66.	GN	Gene Name
67.	GP	Group Authors
68.	GS	Geospatial
69.	GT	Time
70.	HC	Highly Cited Status
71.	HO	Conference Host
72.	HP	Hot Paper Status
73.	ID	Keywords; Identifying Codes
74.	IO	Issuing Organization
75.	IP	International Patent Classification
76.	IS	Issue
77.	IV	Investigators
78.	J9	Journal Abbreviation
79.	JC	NLM Unique ID
80.	JJ	Journal ISO Abbreviation
81.	LA	Language
82.	LS	Language of Summary
83.	LT	Literature Type
84.	MA	Meeting Abstract
85.	MC	Major Concepts or Derwent Manual Code(s)
86.	ME	Medium
87.	MH	MeSH Terms
88.	MI	Miscellaneous Descriptors
89.	MN	Markush Number
90.	MQ	Methods & Equipment
91.	NM	Personal Name Subject
92.	NO	Comments, Corrections, Erratum

93.	NP	Named Person
94.	NR	Cited Reference Count
95.	NT	Notes
96.	OA	Open Access Designations
97.	OB	Record Owner
98.	OC	Country of Original Patent Application Number
99.	OD	Method
100.	OI	ORCID numbers
101.	OP	Original Patent Application Number
102.	OR	Organism Descriptors; Systematics
103.	OS	Original Source
104.	P1	Part Number
105.	P2	Chapter Count
106.	PA	Publisher Address
107.	PC	Country of Patent
108.	PD	Publication Date; Patent Details
109.	PE	Published Electronically
110.	PG	Number of Pages
111.	PI	Publisher City; Patent Priority Information
112.	PM	PubMedID
113.	PN	Part Number; Patent Number
114.	PR	Parts, Structures & Systems; Price
115.	PS	Pages
116.	PT	Publication Type
117.	PU	Publisher
118.	PV	Place of Publication
119.	PY	Publication Year
120.	RC	Date Created, Date Completed, Date Revised
121.	RG	Derwent Registry Number
122.	RI	ResearcherIDs; Ring Index Number
123.	RP	Reprint Address
124.	S1	Source Title (non-English)

125.	SA	Status
126.	SC	Research Areas
127.	SD	Molecular Sequence Data
128.	SE	Book Series Title; Series
129.	SF	Space Flight Mission
130.	SI	Special Issue
131.	SN	ISSN
132.	SO	Source Title
133.	SP	Conference Sponsors
134.	SS	FSTA Section/Subsection; Citation Subset
135.	ST	Super Taxa
136.	SU	Supplement; Research Area
137.	TA	Taxonomic Data
138.	TC	Times Cited Count
139.	TF	Technology Focus Abstract
140.	TI	Article Title
141.	TL	Country of Translation
142.	TM	Geologic Time Data
143.	TN	Taxa Notes
144.	TR	Translators
145.	TS	Translated Source
146.	U1	180 Day Usage Count
147.	U2	Since 2013 Usage Count
148.	UC	Document Selection URL
149.	UR	URL
150.	UT	Accession Number
151.	VL	Volume
152.	VN	Version
153.	VR	Version Number
154.	WC	Web of Science Subject Categories
155.	WE	Web of Science Index
156.	WP	Publisher Web Address

157.	X1	Article Title (non-English)
158.	X2	Article Title (Transliterated)
159.	X4	Spanish Abstract
160.	X5	Spanish Author Keywords
161.	Y1	Portuguese Document Title
162.	Y4	Portuguese Abstract
163.	Y5	Author Keywords (non-English); Portuguese Author Keywords
164.	Z1	Article Title (Other Languages)
165.	Z2	Authors (non-English)
166.	Z3	Publication Name (Chinese)
167.	Z4	Abstract (non-English)
168.	Z5	Author Keywords (non-English)
169.	Z6	Author Address (non-English)
170.	Z7	E-mail Address (non-English)
171.	Z8	CSCD Times Cited Count
172.	Z9	Times Cited, All Databases
173.	ZK	Author Keywords (Korean)

NYILATKOZAT

a diplomadolgozat nyilvános hozzáféréséről és eredetiségéről

A hallgató neve:	Dr. Szalay Zsigmond Gábor
A Hallgató Neptun kódja:	NSN7Z2
A dolgozat címe:	Bibliográfiai adatok elemzésének összehasonlító vizsgálata különböző adatvizualizációs szoftverekkel
A megjelenés éve:	2024
A konzulens intézetének neve:	Műszaki Intézet
A konzulens tanszékének a neve:	Mérnökinformatikai Tanszék

Kijelentem, hogy az általam benyújtott diplomadolgozat egyéni, eredeti jellegű, saját szellemi alkotásom. Azon részeket, melyeket más szerzők munkájából vettem át, egyértelműen megjelöltem, és az irodalomjegyzékben szerepeltettem.

Ha a fenti nyilatkozattal valótlan állítottam, tudomásul veszem, hogy a záróvizsga-bizottság a záróvizsgából kizár és a záróvizsgát csak új dolgozat készítése után tehetek.

A leadott dolgozat, mely PDF dokumentum, szerkesztését nem, megtekintését és nyomtatását engedélyezem.

Tudomásul veszem, hogy az általam készített dolgozatra, mint szellemi alkotás felhasználására, hasznosítására a Magyar Agrár- és Élettudományi Egyetem mindenkori szellemi tulajdon-kezelési szabályzatában megfogalmazottak érvényesek.

Tudomásul veszem, hogy dolgozatom elektronikus változata feltöltésre kerül a Magyar Agrár- és Élettudományi Egyetem MATER Hallgatói Dolgozatok repozitóriumába. Tudomásul veszem, hogy a megvédett és

- nem titkosított dolgozat a védést követően
- titkosításra engedélyezett dolgozat a benyújtásától számított 5 év eltelté után

nyilvánosan elérhető és kereshető lesz az Egyetem MATER Hallgatói Dolgozatok repozitóriumában.

Kelt: Gödöllő, 2024. április 21.



Hallgató aláírása

NYILATKOZAT

Dr. Szalay Zsigmond Gábor hallgató Neptun azonosítója: NSN7Z2 konzulenseként nyilatkozom arról, hogy a záródolgozatot/szakdolgozatot/diplomadolgozatot/portfóliót¹ áttekintettem, a hallgatót az irodalmi források korrekt kezelésének követelményeiről, jogi és etikai szabályairól tájékoztattam.

A záródolgozatot/szakdolgozatot/diplomadolgozatot/portfóliót a záróvizsgán történő védésre javaslom / nem javaslom².

A dolgozat állam- vagy szolgálati titkot tartalmaz: igen nem^{*3}

Kelt: Gödöllő, 2024. év április hó 21 nap


belső konzulens