

DANYI Melinda

Látható, kereshető és megtalálható tudományos publikációk

Akadémiai keresőoptimalizálás (ASEO)

Bevezetés

1982-ben John Naisbitt jövőkutató írta Megatrends című könyvében: „*we are drowning in information but starved for knowledge*” („fuldoklunk az információban, de éhezünk a tudásra”). Bár azóta több, mint negyven év eltelt, ez a megállapítás napjainkban aktuálisabb, mint valaha (Naisbitt, 1982). A digitális tartalmak és információk száma szinte exponenciálisan növekszik a tudományos platformokon, archívumokban és adatbázisokban egyaránt (Baker and Powell, 2024).¹ Ugyanakkor nem csak a publikációk száma nő, de azok megjelenési formája és technológiai háttere is egyre nagyobb változatosságot mutat (új kimeneti formátumok, adatbázisok, forráskódok, preprintek stb.). Ebben a hatalmas információs útvessztőben a tudományos munkák nagy százaléka (tudományágtól függően a kutatási cikkek 7–38%) láthatatlan marad (Kraker et al., 2021). Ez a válság a tudomány hatékonyságára, eredményességére és minőségére egyaránt negatív hatással van.

Az információs túlterheltség nagyban megnehezíti a szerzők számára, hogy láthatóvá tegyék saját munkájukat, és ezzel párhuzamosan a kutatókat és olvasókat is nehéz feladat elé állítja a hatalmas információáradatban való eligazodásban. Ez az ún. felfedezhetőségi válság ugyan nem újkeletű, de az ellene vívott harcra egyre több hadszíntéren szükséges készen állnunk. A tudományos kutatási eredményekre való keresés ma már túlnyomórészt online történik, ezért a siker érdekében elengedhetetlen, hogy azok ne csak ez emberek, de a különböző online keresőmotorok számára is minél egyszerűbben kereshetővé és megtalálhatóvá váljanak.

A tudományos kommunikáció

A tudományos kommunikáció egy ernyőfogalom, mely a tudományos munkák és kutatások folyamatát, azok eredményeinek ellenőrzését, hiteles terjesztését, valamint hosszú távú megőrzését foglalja magában. Napjainkban ennek a folyamatnak hangsúlyos elemévé vált a tudomány népszerűsítése, minél szélesebb körű láthatóvá tétele, melyet a társadalmi és technológiai változások és a megváltozott társadalmi igények (tudomány mediatizálódása) hívtak életre. Jelen cikk a tudományos folyóiratokban való publikáláshoz nyújt olyan információkat és ajánlásokat, melyek hozzájárulhatnak ahhoz, hogy egy kéziratból, látható, kereshető megtalálható tanulmány szülessen.

IA, SEO és ASEO

A tudományos kommunikáció alappillére az Információs Architektúra (Information Architecture, IA), vagy információépítéset, amely az információk megtalálhatóvá és érthetővé tételének tudományként definiálható.² Magába foglalja a keresést, a böngészést, a releváns és kontextuális információk kategorizálását és bemutatását, az ezt biztosító informatikai infrastruktúra biztosítását, azaz a láthatóság, kereshetőség, megtalálhatóság alapfeltételeit. A fentebb taglalt információs túlterheltség mellett a tudományos publikálás és kutatásértékelés követelményeit is szem előtt tartva a tudományos publikálás területén megkerülhetetlen folyamatról és tudományterületről beszélünk (Koltay, 2014).

Ennek egyik eszköze a keresőoptimalizálás (Search Engine Optimization, SEO) ami az online és digitális tartalmak (weboldalak, blogok, közösségi média felületek, géppel olvasható digitális fájlformátumok) keresőmotorok általi organikus megtalálhatóságának javítását, elősegítését jelenti. Organikusnak nevezzük azokat a találatokat, amelyek indexelése és megjelenése nem fizetett szolgáltatás eredménye, azaz tartalmi minőségük alapján, természetes úton kerültek be egy indexbe. A keresőmotorok az ilyen találatokat azok metaadat-szinten értékes tartalma alapján válogatják és rangsorolják.

¹<https://theconversation.com/scientists-around-the-world-report-millions-of-new-discoveries-every-year-but-this-explosive-research-growth-wasnt-what-experts-predicted-237274>

²<https://www.interaction-design.org/literature/topics/information-architecture>

Az elsősorban az online marketing területén használt SEO-stratégia a kereskedelem mellett számos más területen is jól alkalmazható. A keresőoptimalizálás megcélozhat olyan specifikus tartalmakat is, mint a tudományos publikálások és a tudománykommunikáció egyéb platformjai (folyóiratok, könyvek, konferenciaközlmények, blogok, interaktív PDF-tartalmak, online publikációs felületek stb.). Az akadémiai keresőoptimalizálásnak (Academic Search Engine Optimization, ASEO) (Beel et al., 2010; Schilhan et al., 2021) meghatározó szerepe van a tudományos tartalmak létrehozásának, publikálásának, hosszú távú megőrzésének folyamatában, és nagyban segíti a keresett tartalom tudományos adatbázisokban és rangsorokban való megtalálhatóságát oly módon, hogy azokat az akadémiai keresőmotorok számára láthatóvá, feldolgozhatóvá és értékelhetővé teszi.³

De míg egy kereskedelmi weboldalt elég a főbb, piacvezető keresők (Google, Bing, Amazon) számára optimalizálni, addig a tudományos tartalmakat számos különböző keresőnek szükséges lehet megfeleltetni, mert – bár vannak prioritizált, nemzetközi adatbázisok (pl.: Google Scholar, Scopus, Web of Science stb.), de a különböző tudományterületeken, földrajzi régiókban, de akár intézményi szinten is más-más adatbázisok számítanak preferáltnak, ami rendkívül bonyolulttá teszi a keresőoptimalizálás folyamatát.

Tovább nehezíti a kutatók és a tudományos kiadók dolgát az, hogy míg egy üzleti weboldal tartalma szabadon módosítható (hivatkozások szabad módosítása, hozzáadása vagy törlése, a szöveg átírása) – sőt ez a kereskedelemben elvárás is a siker érdekében – és így módon könnyen alkalmazkodhatnak a különböző rangsoroló algoritmusokhoz, addig akadémiai platformok ezt nem tehetik meg. A tudományos hitelesség biztosítása alapvető érdek és cél, így, ha egy tanulmány megjelenik, annak utólagos módosítása nem megengedett. Ezért fontos az ASEO-t különösen átgondoltan és legalább alapfokú SEO-ismereteket alkalmazva végrehajtani. (Schilhan et al., 2021) Miután a kutatásértékelés folyamatában még mindig a publikációk számának, megtekintéseinek, letöltéseinek és hivatkozásainak mátrixa a kutatók és a kutatás elsődleges tudományometriai mércéje, nem szükséges hangsúlyozni, hogy mennyire fontos a jó tartalom, a láthatóság és a kereshetőség.

Az ASEO alkalmazása során rendkívül fontos az etikus magatartási normák betartása. Míg a kereskedelmi célú SEO megengedheti magának a túlzást és „szépítést” a jobb láthatóság érdekében, addig egy tudományos tartalomnál ez nem alkalmazható. Meg kell találni az egyensúlyt az alkotói szabadság, a publikációs kultúra, a kutatási integritás, a láthatóság és a megtalálhatóság között (Sengupta, 2017).

A keresőmotorok működése

A keresőrendszerek többsége mesterségesintelligencia-támogatású, céljuk nemcsak a keresőkifejezések megtalálása, hanem a találatok relevancia szerinti rangsorolása is. (Schilhan et al., 2021) A relevancia-rangsor metódusának meghatározó kulcselemei a bibliográfiai metaadatok. A robotok olyan információkat (metaadatokat) elemeznek és rangsorolnak a szövegben, mint a cím, alcím, a szerző neve, dokumentum- és szerzőazonosítók és kulcsszavak. A keresőrendszer ezen metaadatok tekintetében méri fel a keresett kifejezés gyakoriságát, és pozícióját elsősorban a címbe, az absztraktban és a kulcsszavakban. A nyílt hozzáférésű (open access) publikációkban természetesen a teljes szövegben történik a keresés.

Általánosságban elmondható, hogy minél gyakrabban jelenik meg egy keresési kifejezés a szövegben, az algoritmus annál fontosabbnak fogja rangsorolni az adott tartalmat. A metaadatok – attól függően, hogy hol helyezkednek el a dokumentumban – lehetnek értékesebbek és kevésbé relevánsak. A legfontosabbnak értékelt kulcsszavakat és kifejezéseket minden esetben a címbe helyezzük el.

A keresőmotorok – technológiától függően – egyéb más adatokat is figyelembe vehetnek, úgymint a megjelenés éve, az idézetek és nézettségek a talált dokumentumok teljes számához viszonyítva, valamint a befogadó folyóirat mérőszáma (pl. WoS – impakt faktor, SCImago – Q-besorolás). Ezek közül több tényezőre is befolyásunk lehet csupán azzal, hogy hogyan készítjük el publikációnk keresőbarát PDF-dokumentumát.

³<https://openjournalsystems.com/academic-search-engine-optimization/>

Az ASEO folyamata

A következőkben az akadémiai keresőoptimalizálás legfontosabb lépéseit veszem végig, útmutatást adva a tudományos tartalmak láthatóbbá és kereshetőbbé tételéhez, illetve felhívva a figyelmet a kerülendő hibákra. Előbb a **metaadatoláshoz** szükséges ismereteket veszem végig, majd ezek alkalmazását a tanulmány technikai (PDF) **szerkesztése** során.

Metaadatok használata

Kulcsszavak

Tanulmányunk keresőre optimalizálásakor több eszköz is segítségünkre lehet. Azonban, mielőtt ezeket az online vagy nyomtatott segédleteket használatba vennénk, érdemes a saját tapasztalatainkat is felidézni, összegyűjteni, és átgondolni, hogy mi hogyan keressünk egy adott tartalomra, illetve korábbi kutatásaink során mely kifejezéseket tapasztaltuk a legalkalmasabbnak, melyikkel találkoztunk leggyakrabban. Ezek után a különböző tudományterületi teauruszokban, a szinonimaszótárakban vagy akár a Google keresőszó-generátorában a Google Trend-ben (bár ez utóbbi talán a legkevésbé alkalmas tudományos és szakszövegek optimalizálására) tudjuk ellenőrizni az általunk jónak gondolt kifejezések létjogosultságát, illetve plusz szavakat gyűjthetünk segítségükkel.

Fontos, hogy a kulcsszavakat lehetőleg **egyes számban** használjuk, mert így a ragozott és toldalékolt alakokra is rábukkanhatunk, azok szótöveként. Például a *könyvtár* kulcsszó egyaránt eredményesen használható a *könyvtáros*, *könyvtári*, *könyvtárak*, *könyvtárban* stb. szavakat tartalmazó címek esetében is.

Használhatunk továbbá **általánosabb és specifikusabb kifejezéseket** metaadatként. Az előbbinél nagyobb lesz a merítés, de valószínűleg kevesebb releváns találatot fog tartalmazni a találati lista, míg egy szakspecifikus kifejezésnél kevesebb, de használhatóbb találatra számíthatunk. Például a *Magyar templomépítéset a középkorban* című tanulmány esetén az *építéset* kulcsszó önmagában nem fogja a munkánk megtalálását elősegíteni. Célravezetőbb például a *magyar építéset*, *templomépítéset*, *középkori építéset* kifejezéseket, szókapcsolatokat használni.

Fontos figyelni arra is, hogy mennyi kulcsszót használunk egy dokumentumon belül. Számuk tekintetében nincsen nemzetközileg általánosan elfogadott szabály, csupán ajánlások⁴ melyek **három–nyolc kulcsszó** használata javasolják. Ezt nem érdemes túllépni, hisz a feleslegesnek értékelt kulcsszavakat a keresőmotorok automatikusan letiltják a találati listából, azokat ún. „kulcsszó szemét”-ként (keyword spam) minősítik⁵. Magyarán nem lesz sikeresebb a publikációnk, ha több keresőszót adunk meg hozzá, sőt a túl sok kulcsszó használata rontja is annak megtalálhatóságát.

A nagyon **ritka, egyedi szavak** esetén érdemes annak egy gyakrabban használt szinonimáját, vagy magyarázó kifejezését is felvenni a kulcsszavak közé (és persze a szövegtörzsben is használni ezeket).

A rendszeresen publikáló szerzőknek érdemes lehet saját kulcsszókészletet összeállítani, amiből később könnyen meríthet, megspórolva ezzel a kulcsszókutatásra szánt időt.

A cím és alcím optimalizálása

A tudományos tartalmak legrelevánsabb eleme és egyben metaadata a cím. A keresőmotorok számára a címben előforduló kulcsszavak és keresőkifejezések bírnak a legnagyobb relevanciával. Itt érdemes megjegyezni, hogy nem csak az algoritmusok, de a humán felhasználás szempontjából nagyon fontos, hogy a legfontosabb kifejezések azonnal azonosíthatók legyenek a címben, hisz egy olvasó ezzel az információval találkozik legelőször.

⁴Ajánlott számuk folyóiratokként eltérő lehet. A nagy nemzetközi lapok általában 3-7, 4-8 vagy 5-7 darabszámban limitálják a használatukat.

⁵<https://www.seo.com/basics/glossary/keyword-stuffing/> ; <https://foxdata.com/en/glossary/keyword-spamming/>

A cím megalkotásának első szabálya: legyen rövid és lényegre törő. Kutatások kimutatták, hogy a hosszabb címek általában kevesebb hivatkozást kapnak, mint a rövid, tömör címek, (Letchford et al., 2015) ami azért fordulhat elő, mert egyes keresőmotorok, mint pl. a Google, csak a cím első 65 karakterét jelenítik meg, így a hosszú címek végén szereplő kulcsszavak haszontalanná válnak (Montti, 2021). Ugyanez a helyzet a mobileszközökön való olvasással, ahol a nem reszponzív tartalmak csonkán vagy átláthatatlanul jelennek meg. Ebből következik, hogy a legfontosabb kulcsszavakat a cím elején, annak lehetőleg az első 65 karakterében szerepeltessük.

Sok esetben egy tanulmányt a főcím mellett **alcímmel** is ellátnak, ami szintén buktatókat hordozhat. Az egyik, látszólag kreatív és játékos címadási módszer az, hogy a főcím egy idézetet, szójátékot tartalmaz, vagy egyéb kreatív, asszociatív szöveget, és a tanulmány valós, tudományos terminusokkal való leírása csak az alcímben szerepel.

Pl.: Főcím: „Vár állott, most kőhalom...”, alcím: A magyar végvárrendszer építészeti öröksége

Míg a főcímbe maximum egy kulcsszónak alkalmas kifejezést találunk (vár), addig az alcím hatékonyan optimalizálható: magyar vár, végvárrendszer, építészet, építészeti örökség

Egy ilyen cím az olvasókban pozitív benyomást kelthet és az ember számára egyértelmű tartalmi kapcsolatokat jelenthet, de a robotok számára kevésbé vagy hibásan (pl. költészet tematikába sorolás – Himnusz) értelmezhető. Gépi kereséskor az alcím kevésbé releváns metaadat (vannak adatbázisok, amik nem is jelenítik meg az alcímet a találati listában), így hiába tartalmaz valid keresőszókat vagy kifejezéseket, a tanulmány érdemtelenül alulértékelt lehet, a találati lista sokadik helyére sorolódhat.

Hasonló a helyzet a **különleges karakterek** esetében is (kiskötőjel, nagyköötőjel, gondolatjel, mínuszjel, alulvonás, idézőjel, csillag, perjel stb.). Míg egy olvasó gond nélkül értelmezi a rövid- vagy hosszú kötőjelekkel ellátott kifejezések szemantikai kapcsolatát, vagy egy idézőjel jelentését egy címben, addig egy keresőrobotnak ez problémát okozhat. Például a *közép- és nagyvállalati innováció* kifejezés esetén egy ember számára egyértelmű, hogy közép- és nagyvállalati fejlesztésekről szól a tanulmány, de egy keresőrobot ezt nem „érti”. Arról nem is beszélve, hogy pl. a Google Scholar az ilyen karaktereket tartalmazó címeket sok esetben hibásan – szóköz nélkül – jeleníti meg: *közép-és nagyvállalati*, és ezzel már el is jutottunk a különleges karakterek használatának következő hátrányához: a **hibás idézéshez**. Minél több speciális karaktert tartalmaz egy cím, annál nagyobb a valószínűsége, hogy azt hibásan fogják idézni (főleg, ha maga az adatbázis is hibásan jeleníti meg). (1. ábra) Ezt úgy tudjuk elkerülni, ha mindkét fogalmat külön-külön kiírjuk: *középvállalati és nagyvállalati innováció*, amely kifejezésben mindhárom szót kiemelten alkalmas kulcsszóként is feltüntetni: *középvállalati, nagyvállalati, innováció*.

1. ábra — A Google Scholar találati listájában szereplő kötőjeles kifejezések hibás megjelenítése

IDÉZET] A termékenység néhány jellemzője **Közép- és Dél-Európában az első világháború előtt**

T Kálmán - Demográfia, 1969 - demografia.hu

A demográfiai tranzíció elmélete—azon demográfiai változások általánosításai, amelyek során a népesség korábbi magas termékenysége és halandósága jelentős csökkenésen ment ...

☆ Mentés Hivatkozás Idézetek száma: 18 Kapcsolódó cikkek Mind a(z) 2 változat

Magyar **közép- és nagyvárosok versenyképességi tipológiája és humán tőké**
vonzó potenciálja

V Varga, D Teveli-Horváth, G Horzsa - Földrajzi Közlemények, 2021 - ojs.mtak.hu

A globalizációs versenyképesség szempontjából a régiókkal szemben egyre nagyobb szerepet kapnak a városok és a városrészek, illetve ezzel párhuzamosan a város-vidék ...

☆ Mentés Hivatkozás Idézetek száma: 5 Kapcsolódó cikkek Mind a(z) 8 változat

Milyen tényezők akadályozzák Magyarországon a **közép- és nagyvállalati innovációt?**

J Kiss - Vezetéstudomány/Budapest Management ..., 2021 - journals.lib.uni-corvinus.hu

The author analysed the barriers to the innovation activity of companies based on the latest survey of the BCE Competitiveness Research Centre. He found that four of the twelve ...

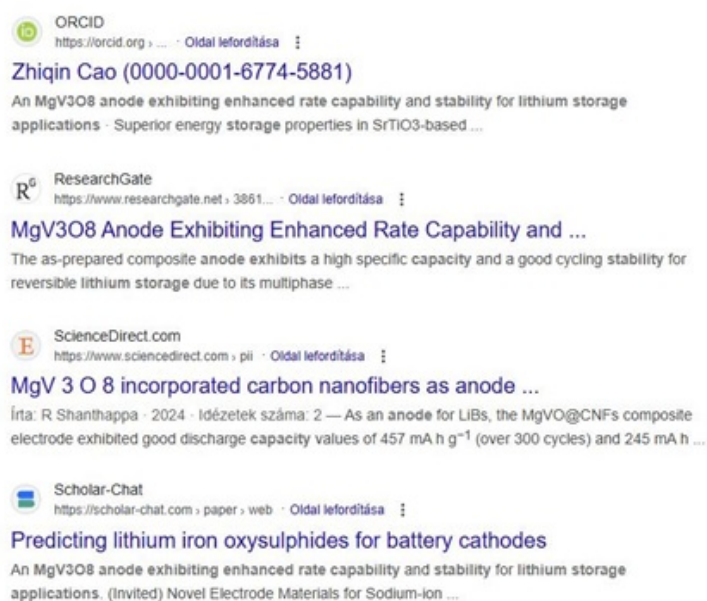
☆ Mentés Hivatkozás Idézetek száma: 6 Kapcsolódó cikkek Mind a(z) 5 változat

A **felső- és alsó indexek** szintén okozhatnak megjelenítési és idézési problémát. Ahogy a 2. ábrán látható az *An MgV₃O₈ anode exhibiting enhanced rate capability and stability for lithium storage applications* közlemény címében szereplő kémiai kifejezés egyik találat esetén sem jelenik meg helyesen.

Hasonló a helyzet a címbe ágyazott **egyenletek** esetén, ha nincs megfelelő karakterkódolás (UTF-8.11) technikailag biztosítva. A 3. ábrán látható, hogy a Google teljesen átkódolja, így az szinte felismerhetetlen a találati listában, a nagy adatbázisok (WoS, Scopus) pedig error-jelzést adhatnak, ha az ilyen címet bemásolva indítunk keresést.

A karakterkódolásra akkor is fontos hangsúlyt fektetni, ha görög, cirill, vagy egyéb, az alap latin karakterkészlethől eltérő karaktert (λ , ψ , ϕ , Ω , Π , $\ddot{a}$, å , æ , $\emptyset$) tartalmaz a szöveg és a cím.

2. ábra — Alsó- és felső indexek hibás megjelenítése a Google találati listában



3. ábra — Az egyenletek hibás megjelenése a címben

A tanulmány címe az eredeti PDF/A-ban

ON THE EQUATION

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u) + \lambda|u|^{p-2}u = 0$$

Találat a Google-ban

On the Equation $\operatorname{div}(|\nabla u|^{p-2}\nabla u) + \lambda|u|^{p-2}u = 0$
 Írta: P. Lindqvist · 1990 · Idézetek száma: 891 — $\operatorname{div}(|\nabla u|^{p-2}\nabla u) + \lambda|u|^{p-2}u = 0$. PETER LINDQVIST.
 (Communicated by Barbara L. Keyfitz). ABSTRACT. The first eigenvalue $\lambda = \lambda_1$ for the equation...

Címre keresés a WoS-ban

Search Error: Missing Boolean operator.

All Fields

+ Add row + Add date range Advanced search

Clear Search

Címre keresés a Scopus-ban

Search within Search documents

+ Add search field

Documents Preprints Patents Secondary documents Research data ?

Your search contains a syntax error

- Make sure there are not special characters in your query
- Check if there repeated operators
- Make sure your operators are in the correct order

Az angol nyelvű tartalmak esetén érdemes megfontolni, hogy melyik angol nyelvjárást használjuk (leggyakoribb a brit vagy amerikai). Ennek okát legjobban úgy lehet szemléltetni, ha egy angol szóra mindkét változatban rákeresünk. A Google Scholar-ban például a *digitalizál* szó esetén brit helyesírással (digitise) ~32.600, amerikai helyesírással (digitize) ~324.000 találatot kapunk.⁶

Végül érdemes még szót ejteni a **rövidítések és mozaikszók** használatáról, amik szintén ronthatják a kereshetőséget. Ezeket általában érdemes csak abban az esetben használni, ha az jól ismert és gyakran használt az adott tudományos területen, illetve javasolt a feloldásukat is feltüntetni a szövegben.

Az absztrakt optimalizálása

A cím után az absztrakt a tudományos közlemények második legfontosabb tartalmi eleme.⁵ Optimális terjedelme 200–250 karakter. A keresőmotorok számára az itt megfogalmazott szöveg a cím után a következő legrelevánsabb tartalom, és nem kevésbé fontos, hogy a cikk szakmai bírálói és az adatbázisokban kutatók is ezzel a tartalommal találkoznak rögtön a cím elolvasása után. A legtöbb esetben az absztrakton múlhat a dolgozat további sorsa, ezért ennek szakszerű megírása ugyanannyira fontos, mint a címé. Amellett, hogy a kutatómunkát, annak módszereit, eredményeit és a konklúziókat figyelemfelkeltően, ugyanakkor tömören kell tartalmaznia, a címhez hasonlóan a megfelelő kulcsszavakat és -kifejezéseket is tartalmaznia kell. Törekedni kell arra, hogy az első egy-két mondatban előforduljanak a legfontosabb szavak, kifejezések, mivel nem minden adatbázis és keresőprogram jeleníti meg a teljes absztraktot, csak annak első egy-két mondatát (első 160 karakter) (Cushman, 2018).⁷

Az absztrakt megírásakor hasonló módon kell eljárni, mint a cím esetében, vagyis törekedni kell az általános és specifikus szavak egyensúlyára, megfelelő szinonimák alkalmazására, de ezeket a szavakat már érdemes a tanulmány további részeiben is következetesen alkalmazni, hogy a keresőmotorok azt valóban releváns kifejezésként értelmezzék. A kulcsszavakat itt érdemes lehetőség szerint (a terjedelemmel arányosan és a szöveg minőségét nem rontva) többször is szerepeltetni. A túloptimalizálást itt is – ahogy a teljes szövegben – kerülni kell.

⁶2025. január 22-i lekérdezés

⁷<https://www.sfdit.net/abstract-content-optimization-the-key-to-unlocking-journal-acceptance/>

Az absztraktra vonatkozó tartalmi és formai előírást a legtöbb folyóirat közli a szerzői útmutatóban, ami tartalmazhat egyéb követelményeket is, úgymint terjedelmi korlát szó- vagy karakterszámban megadva, vagy tartalmi struktúra. Ezeknek minden esetben meg kell felelni, a keresőre optimalizálás ezeket az elvárásokat nem sértheti.

Felhasznált irodalom és hivatkozások

Ahogy a kereskedelmi célú tartalomoptimalizálásnál, úgy a tudományos keresőoptimalizálásnál is elengedhetetlen a megfelelő linképítés. Ennek alapszabálya, hogy minden, a tanulmányban előforduló linket – legyen az a nyitó oldalon lévő azonosító (DOI, ORCID) vagy az irodalomjegyzékben lévő hivatkozás – url-lel kell ellátni. Ez nem csak azért szükséges, hogy rákattintva elnavigálja a kutatót a megfelelő tartalomhoz, de a keresőmotorok is ezek segítségével aratják az adatokat, melyek mintegy átjárókat biztosítanak számukra a különböző tartalmak, adatbázisok és weboldalak között. Minél több működő linket tartalmaz a cikkünk, annál több keresőmotor fogja elérni azt, ezáltal gyarapíthatjuk a bejövő hivatkozásaink számát. Ez a fajta „cserekereskedelem” természetesen csak akkor működik, ha azok a publikációk is el vannak látva működő url-ekkel, amire hivatkozunk, és amelyik hivatkozik ránk.

Állandó azonosítók használata

A metaadatok egyik széles körben elterjedt formája az állandó azonosítók, melyek szolgálhatnak a dokumentum azonosítására (DOI, Digital Object Identifier, magyarul digitálisobjektum-azonosító), személyek (ORCID, Open Researcher and Contributor ID) vagy a kutatást végző intézmények (Research Organization Registry, ROR) identifikálására (Holl és Bilicsi, 2023). Elhelyezésük a cikk nyitó oldalán, jól látható helyen ajánlott: a DOI-t a fejléc vagy láblécbe, míg az ORCID-et a szerzői affiliációba a legoptimálisabb szerkeszteni.

A DOI akkor is képes beazonosítani egy tanulmányt, ha annak URL-je, elérési útvonala idővel megváltozik vagy megszűnik. Használata a nagyobb nemzeti és nemzetközi kiadóknál már kötelező, az irányadó tudományos adatbázisok is alapkövetelménynek tekintik, és egyre több kisebb folyóirat is felismeri fontosságát. Az ORCID többek között az azonos nevű szerzők (pl.: gyakori vezetéknév: Kovács, Tóth, Nagy) megkülönböztetésében és névváltozások esetén (házastárs nevének felvétele) nyújt nagy segítséget. A publikáció valós szerzőjéhez való hozzárendelése kulcsfontosságú nemcsak a kutatók tudományos pályafutása, de az intézmények kutatási eredményeinek mérése szempontjából is.

Nem lehet eléggé hangsúlyozni, hogy ezek feltüntetése minden esetben URL-lel ellátva, azaz „kattintható” állapotban történjen. Ha csak sima szöveggént jelennek meg PDF-ben, az a keresőrobotok számára láthatatlan.

Creative Commons licenz

Az Open Access tartalmak nagyon fontos adatai a felhasználására vonatkozó információk, melyeket a Creative Commons (CC) licenzek rögzítenek. Amellett, hogy ez a nyílt hozzáférés lehetővé teszi, hogy a tudás ne szoruljon fizetős korlátok közé és szabadon elérhető legyen a nagyközönség számára, a szerzők a különböző CC-licenzek megadásával lehetőséget biztosítanak a publikáció felhasználására és terjesztésére úgy, hogy közben szerzői jogaik nem sérülnek. Emellett a licenz megadása megkönnyíti az adatbázisok közötti adat- és információcserét, a tudományos eredmények disszeminációját. A leggyakoribb Creative Commons konstrukciókat az **1. sz. melléklet** mutatja be.

A keresőbarát PDF/A-tartalmak

Ebben a fejezetben azt mutatom be, hogy miként kell elkészíteni egy PDF/A-formátumú tudományos tanulmányt, hogy az megfeleljen a láthatóság, kereshetőség és megtalálhatóság technikai feltételeinek. Ennek grafikai megjelenítése az **1. számú mellékletben** látható.

A tanulmány elengedhetetlen metaadatai

Ahhoz, hogy egy tanulmány PDF-dokumentuma szakmailag és technikailag megfeleljen a publikációs követelményeknek és a metaadat-szintű adatcseré megvalósulásának az alábbi elemeket szükséges tartalmaznia (a tanulmány törzsszövegén kívül):

- cím (és alcím)
- szerző(k) neve(i)
- szerző(k) affiliációja/i (legalább az intézmény neve, a levelezőszerző esetén az email cím)
- absztrakt (legalább a tanulmány elsődleges nyelvén)
- kulcsszavak (legalább a tanulmány elsődleges nyelvén)
- DOI
- ORCID
- Creative Commons licenc (grafikusan vagy szövegesen, linkelve)
- folyóirat, lapszám, cikk adatok (folyóirat neve, évfolyam, lapszám, oldalszám vagy cikkazonosító)
- DOI-val és egyéb url-ekkel ellátott hivatkozáslista

Ezek szerkesztésére vonatkozó ajánlást az **2. sz. melléklet** tartalmaz.

Szöveg formázása

A megfelelően strukturált szöveg segít a keresőrobotoknak a tartalom rendszerezésében és azonosításában, a tanulmány fókuszának megértésében. A tanulmányok írásának szakmai-tartalmi felépítését leíró IMRAD-szerkezetet (Introduction/Bevezetés, Materials and Methods/Anyagok és módszerek, Results/Eredmények és Discussion/Vita) (Subachev, 2023) minden kutató ismeri, de ez önmagában még nem jelenti a szöveg alkalmasságát a keresethez és megtalálhatóság szempontjából.

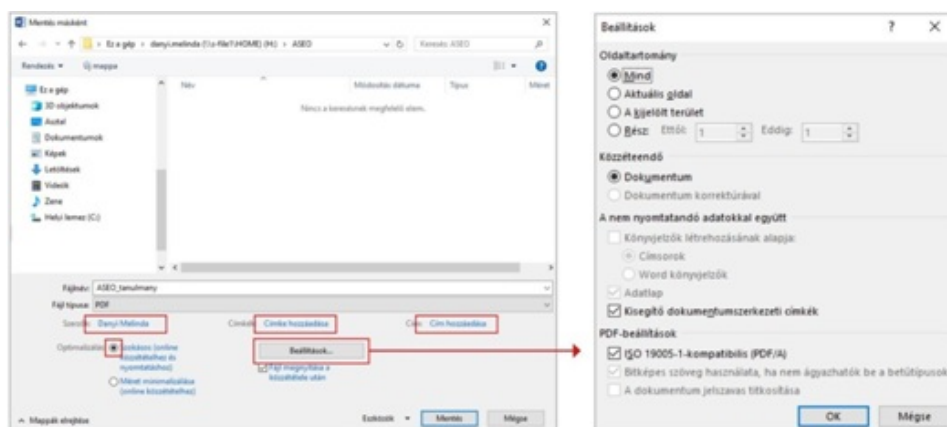
A nyitó oldalon elhelyezett főcím (és alcím), szerzők nevei, affiliációjuk, az absztrakt és a kulcsszavak után a tanulmány teljes tartalmában is szükséges bizonyos szöveghierarchiát felépíteni. Ezek a html-jelölőnyelvben használt <h1></h1>-elemek után *H1- H2- H3-hierarchiának* nevezzük, ami a dolgozat címeinek bizonyos betűméret és -tulajdonság beállításával történő kiemelését jelenti. Magyarul a magasabb hierarchiájú fejezetcímeket jobban kiemeljük, mint az alacsonyabb szintű (alfejezetcím) tartalmakat. A szöveg strukturálását a felsorolások is segítik, melyek jól megszerkesztve szintén hatékonyan vezetik a keresőrobotokat.

PDF-ek metaadat-optimalizálása

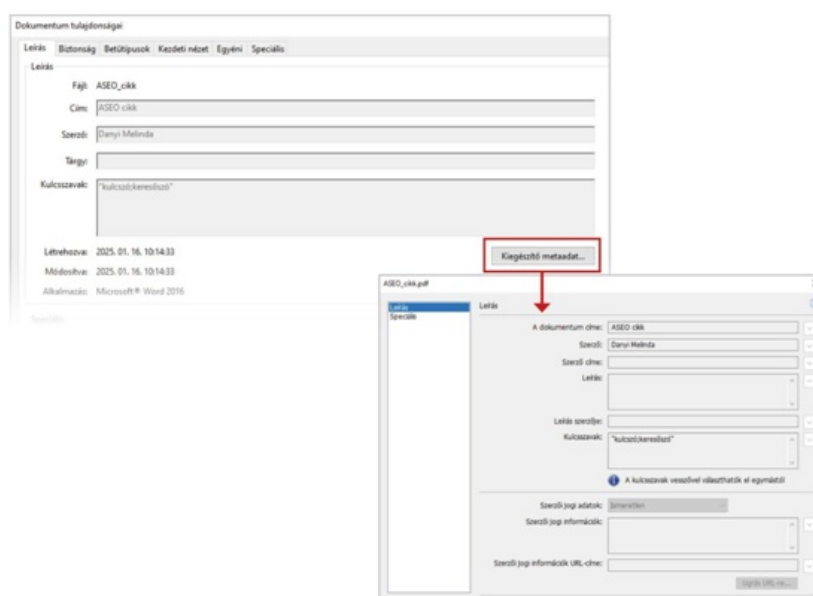
A tudományos tanulmányokat a html-megjelenítés mellett a legtöbb esetben PDF/A-formátumban publikálják a kiadók, amiket megfelelő technikai előállítás esetén a keresőrobotok is olvasni és indexálni tudják. Azonban a szövegszerkesztés és az IMRAD-protokol mellett szükség van metaadat információk hozzáadására is, ahhoz, hogy ez megvalósulhasson. A kiadói változat elkészítésénél az archiválható PDF/A szabványt kell használni, melyet a különböző kiadványszerkesztő szoftverek (Adobe InDesign, Microsoft Word, Open Office, Adobe Acrobat) előállítani és kezelni tudnak. A PDF/A a Portable Document Format (PDF) formátum ISO által szabványosított változata, mely az elektronikus dokumentumok hosszútávú megőrzésére lett kifejlesztve.

Microsoft Word-ben a dokumentum PDF-formátumban való mentésekor van lehetőség megadni a mű szerzőjét, kulcsszavak és a cím hozzáadását (Címke és Cím hozzáadása). A *Beállítások* ablakban az ISO-kompatibilis PDF/A lehetőséget javasolt bejelölni (4. ábra). Később, a már elkészült PDF-ben is van lehetőség a metaadatok szerkesztésére, javítására és kiegészítésére az Adobe Acrobat segítségével (cím, szerző, szerzői információk, kulcsszavak, tárgy, szerzői jogi információk) (5. ábra).

4. ábra — PDF-metaadat szerkesztése MS Word-ben

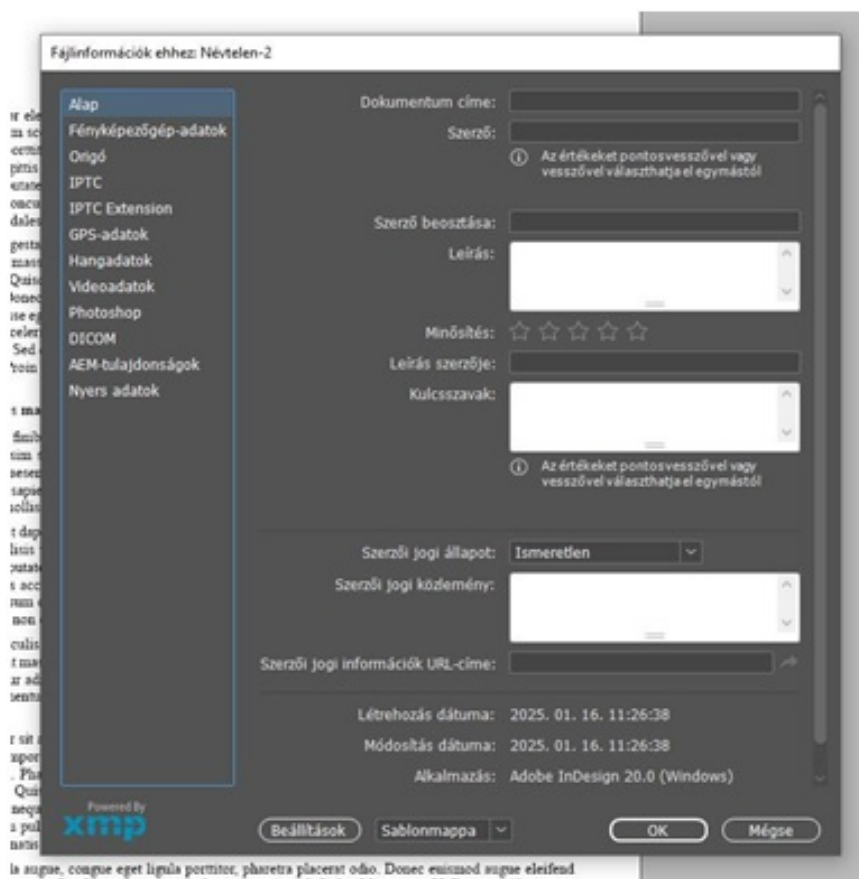


5. ábra — PDF-metaadat szerkesztése Adobe Acrobat-ban



Az *Adobe InDesign*-ban nagyon hasonló a dokumentumok metaadatokkal való ellátása. A *Fájl*-menü alatt a *Fájlinformációk*-ablakba vezethetők fel ugyanazok az információk, mint a MS Word esetén: cím, szerző, szerzői információk, kulcsszavak, tárgy, szerzői jogi információk. (6. ábra)

6. ábra — PDF-metaadat szerkesztése Adobe InDesign-ban



Ábrák és illusztrációk metaadat-optimalizálása

A PDF-formátumú publikációkban külön optimalizálhatók a képek, grafikonok és táblázatok is. Ezek a csatolmányok ún. *alt tag*-el való ellátása hasonlóan történik a MS Word és az Adobe InDesign használatakor. Az előbbiben a képet kijelölve jobb egérgombos kattintással a *Kép formázása* majd az *Elrendezés* és tulajdonságok pont alatt a *Helyettesítő szöveg* résznél lehet megadni a kép leírását kulcsszavakkal de mondatokkal is. InDesign-ban az *Objektum exportálási beállításainál* az *Alternatív szöveg*-fűl alatt lehet ezt megtenni.

Míg az ún. pixelgrafikus fájlok (pl.: jpg, png, bmp stb.) gépi olvasása külön alternatív szöveget (alt tag) igényel, vannak olyan vektorgrafikus képformátumok is, melyekből a keresőmotorok automatikusan ki tudják olvasni a szöveget, ilyen például az svg- és az eps-kiterjesztés. Ezért ezeknél is érdemes figyelni a szövegük megfogalmazására, kulcsszavak beépítésére.

Szövegcentrikus publikálási szabványok

A PDF formátum mellett lehetőség van a tanulmány szövegcentrikus tárolására és megjelenítésére is, ami még optimálisabb forrást jelent a keresőmotorok számára. Az egyik legismertebb ilyen nemzetközi ajánlás a TEI XML, mely kifejezetten a tudományos közösség által és számára készült és szoftveres környezettől függetlenül, formázás nélkül tárolja a szöveget ezzel nagyban megkönnyítve a szövegbányászatot, a gépi feldolgozást és a szemantikus web megvalósulását. A nagy nemzetközi folyóiratoknál már bevett gyakorlat, a hazai platformokon pedig tendencia a kettő – a szöveg PDF/A- és XML-formátumban való együttes publikálása.

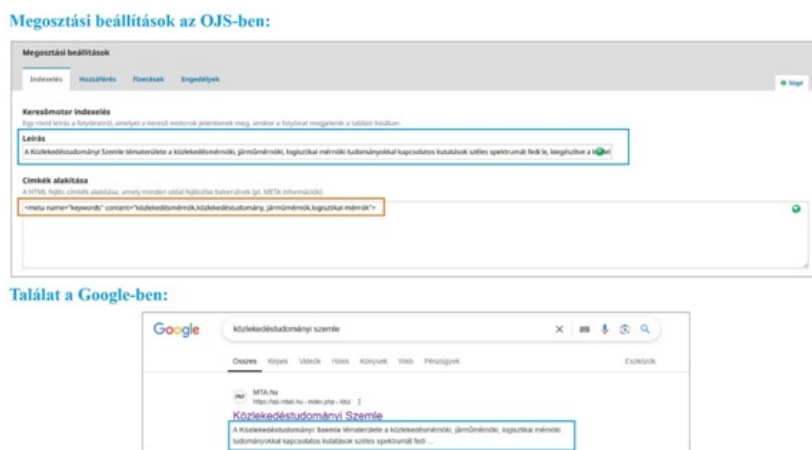
Keresőoptimalizálás az OJS-platformon

Az Open Journal System (OJS) az egyik legnépszerűbb nyílt forráskódú folyóirat-menedzsment platform szabadon hozzáférhető (open access) folyóiratok számára. A publikációk a kéziratok benyújtásától kezdve a cikkek végleges kiadói publikálásáig tartó teljes folyamaton keresztül nyomon követhetők

mind a szerkesztőségi tagok, mind a szerzők számára. A rendszer metaadat szinten több adatbázissal, keresőmotorral és DOI-ügynökséggel is képes kommunikálni (Scopus, Google Analytics, Google Scholar, Crossref, ORCID stb.). A jól működő adatcserének a megfelelő plugin-ok beállítása mellett természetesen az a feltétele, hogy a dokumentum és a szerzők adatai hibátlanul és hiánytalanul legyenek kitöltve az adatmezőkben. Ezek egy részét már benyújtáskor szükséges megadni (pl.: név, cím), van, ami még csak opcionálisan állítható kötelező adatnak ebben a fázisban (pl.: absztrakt, kulcsszó) és bizonyos adatok – értelemszerűen – publikáláskor kerülnek rögzítésre (évfolyam, lapszám, oldalszám stb.)

Az OJS-en belül lehetőség van külön keresőmotorra (pl.: Google, Firefox, Opera) is optimalizálni a folyóiratokat. A megosztási beállításokon belül megadhatjuk a folyóirat jól átgondolt, kulcsszavakkal ellátott rövid leírását, valamint az ezt támogató html-alapú keresőszavakat (<meta name="keywords" content="kulcsszavak">). A 7. ábrán az előbbi a kék-, az utóbbi pedig narancssárga keretben látható. A leírásnál megadott optimalizált szöveg jelenik meg a keresők (az ábrán a Google-keresőben) találati listájában, a html-kód láthatatlan az olvasók számára láthatatlanul, a háttérből támogatja a keresést. Ez a beállítás már haladó informatikai ismereteket igényel, így a szerkesztőségek vagy az informatikus könyvtáros feladata ennek biztosítása. Az ábrán az is jól látszik, hogy a találati listában a leírás legfeljebb első 160 karaktere kerül megjelenítésre, az azon felüli tartalomra csak a hármaspont utal.

7. ábra — Keresőmotorra optimalizálás OJS-ben



Összefoglalás

Az akadémiai keresőoptimalizálás elengedhetetlen ahhoz, hogy a tudományos eredmények láthatók, kereshetők és megtalálhatók legyenek a tudományos adatbázisokban, digitális archívumokban vagy az Interneten. A tudományos irodalom kibocsátása világszerte folyamatosan növekszik, minek következtében a tudóstársadalomnak szembe kell nézniük az információs túlterheltség és a felfedezhetőségi válság jelenségeivel. Segítséget a tudományos kommunikáció alappillére, az Információs Architektúra (IA), vagyis az információk megtalálhatóvá és érthetővé tételének tudománya, azon belül pedig a kulcsszó- vagy metaadat-alapú keresőoptimalizálás jelenthet. Ez a tanulmány útmutatást kíván nyújtani a láthatóság, kereshetőség és megtalálhatóság technikai és informatikai elvárásainak megfelelő publikációk elkészítéséhez.

Felhasznált szakirodalom

- Baker, D. P., Powell, J. J. W., (2024). Scientists around the world report millions of new discoveries every year – but this explosive research growth wasn't what experts predicted, *The Conversation*, <https://theconversation.com/scientists-around-the-world-report-millions-of-new-discoveries-every-year-but-this-explosive-research-growth-wasnt-what-experts-predicted-237274>

- Beel, J., Gipp, B., Eilde, E. (2010). Academic Search Engine Optimization (ASEO) Optimizing Scholarly Literature for Google Scholar & Co., *Journal of Scholarly Publishing*, 41(2), pp. 176–190, <https://doi.org/10.1353/scp.0.0082>
- Cushman, M. (2018). Search engine optimization: What is it and why should we care?, *Research and Practice in Thrombosis and Haemostasis*, 2(2), pp. 180–181, <https://doi.org/10.1002/rth2.12098>
- Holl, A., Bilcsi, E. (2023). Persistent identifiers (PIDs - állandó, egyedi azonosítók) kutatási eredmények azonosítására, <http://real.mtak.hu/id/eprint/183032>
- Jeschke, J. M., Lokatis, S., Bartram, I., Tockner, K. (2019). Knowledge in the Dark, Scientific Challenges and Ways Forward, *FACETS*, 4(1), pp. 423–441, 2019. <https://doi.org/10.1139/facets-2019-0007>
- Koltay, T. (2014). Az információs túlterhelés és az információs műveltség, *Tudományos és Műszaki Tájékoztatás*, 61(3), pp. 83–89.
- Letchford, A., Moat, H. S., Preis, T. (2015). The advantage of short paper titles, *Royal Society Open Science*, 2(8), 150266, <https://doi.org/10.1098/rsos.150266>
- Kraker, P., Schramm, M., Kittel, C. (2021). Discoverability in (a) Crisis, *ABI Technik*, 41(1), pp. 3–12. <https://doi.org/10.1515/abitech-2021-0003>
- Montti, R. (2021). What Is The Best Title Tag Length For SEO?, *Search Engine Journal*, <https://www.searchenginejournal.com/what-is-the-best-title-tag-length-for-seo/427938/>
- Naisbitt, J. (1982). Ten New Directions Transforming Our Lives, *Megatrends* Warner Books, New York ISBN: 0446512516 https://archive.org/details/megatrendstennew0000nais_y7q8/page/n371/mode/2up
- Schilhan, L., Kaier, C., Lackner, K. (2021). Increasing visibility and discoverability of scholarly publications with academic search engine optimization, *Insights: the UKSG journal*, 34(1). <https://doi.org/10.1629/uksg.534>
- Sengupta, S., Honava, S. G. (2017). Publication ethics, *Indian Journal of Ophthalmology*, 65(6), pp. 429–432. https://doi.org/10.4103/ijo.IJO_483_17
- Subachev, Y. (2023). Research Paper IMRAD Structure: Features and Recommendations, English Proofreading Experts. *Introduction, Materials and Methods, Results and Discussion*. <https://english-proofreading-experts.com/imrad-structure/>

Online források

- <https://theconversation.com/scientists-around-the-world-report-millions-of-new-discoveries-every-year-but-this-explosive-research-growth-wasnt-what-experts-predicted-237274>
- https://ixdf.org/literature/topics/information-architecture?srsId=AfmBOoor_RGKA0M6Zvy2oKDE__ImVGdTmgm9CDjysivAoGEr4cSugdH5
- <https://openjournalsystems.com/academic-search-engine-optimization/>
- <https://www.seo.com/basics/glossary/keyword-stuffing/>
- <https://foxdata.com/en/glossary/keyword-spamming/>