

A TUDOMÁNY, AZ OKTATÁS ÉS A KÖZGYŰJTEMÉNYI KISZOLGÁLÁS ÚJ INFORMATIKAI SZINERGIÁI

**NETWORKSHOP 2026
35. Országos Informatikai Konferencia**

**2026. március 31–április 2.
Debreceni Egyetem, Debrecen**

Szerkesztette: Tick József, Kokas Károly, Holl András

**HUNGARNET Egyesület
Budapest, 2026**



HUN-REN
Magyar Kutatási Hálózat

NETWORKSHOP

Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Korrektúra: Danyi Melinda

Angol nyelvi lektor: Lukács Katalin

Networkshop 2026 konferencia előadásainak közleményei

Debreceni Egyetem, Debrecen

2026. március 31–április 2.

ISBN 978-615-6792-29-7

DOI: <https://doi.org/10.31915/NWS.2026>

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével

Budapest

2026

Borítókép: [freepik.com](https://www.freepik.com)

RAG MINT KÖNYVTÁRI TUDÁSRENDSZER

A MESTERSÉGES INTELLIGENCIA HANGJA AZ IR VS. DR VITÁBAN

RAG AS A LIBRARY KNOWLEDGE SYSTEM

THE VOICE OF ARTIFICIAL INTELLIGENCE IN THE IR VS. DR DEBATE

Fülöp Endre

endre.fulop@qulto.eu

Qulto – Monguz, rendszertervező

ELTE BTK Könyvtár- és Információtudományi Intézet, doktorandusz

Absztrakt

Miután Calvin Mooers – bizonyos értelemben Paul Otlet monografikus elvét követve – 1950-ben megalkotja az információ-visszakeresés (Information Retrieval, IR) terminust, a könyvtártudomány szinte azonnal átveszi és magáévá teszi a fogalmat és a mögötte álló elvet. Ezzel lényegében azt az elvárást támasztja a könyvtári tudásrendszerrel szemben, hogy amikor a felhasználó információt keres, akkor a könyvtári rendszer ne pusztán a kérdés szempontjából releváns dokumentumokhoz legyen képes őt irányítani, hanem adjon közvetlenül és azonnal tényszerű választ a megfogalmazott kérdésre. A könyvtári tudásrendszer IR-rendszerként való felfogásának gondolatát nem fogadta egyöntetű lelkesedés. Számos elméletalkotó (többek között Robert Fairthorne, Bernd Frohmann, Birger Hjørland és Rafael Capurro) különböző okokból és nézőpontokból, de egyaránt bírálja a paradigmaváltást, és amellet érvel, hogy a könyvtári tudásrendszer fő feladatának mégis inkább a dokumentum-visszakeresést (Document Retrieval, DR) kellene tekinteni. Tanulmányomban arra a kérdésre keresek választ, hogy a nagy nyelvi modellek (LLM) és különösen a Retrieval Augmented Generation (RAG) megjelenését tekinthetjük-e fordulópontnak ebben az IR vs. DR vitában. Szolgáltat-e a RAG érveket akár az egyik, akár a másik oldal számára? Mit kell, mit lehet, mit célszerű elvárunk egy könyvtári tudásrendszerként működő RAG-alkalmazástól?

Kulcsszavak információfogalom, információfilozófia, információ-visszakeresés, dokumentum-visszakeresés, mesterséges intelligencia, LLM, RAG, metainformáció

Abstract

After Calvin Mooers coined the term Information Retrieval (IR) in 1950 – in a sense following Paul Otlet’s monographic principle – library science almost immediately adopted and integrated both the concept and its underlying principle. This essentially imposed an expectation on library knowledge systems: when a user searches for information, the system should not merely direct them to documents relevant to their query, but should provide a direct and immediate factual answer to the formulated question. The notion of conceiving the library knowledge system as an IR system was not met with unanimous enthusiasm. Numerous theorists (including Robert Fairthorne, Bernd Frohmann, Birger Hjørland, and Rafael Capurro), for various reasons and from different perspectives, criticized this paradigm shift, arguing that the primary task of library knowledge systems should remain Document Retrieval (DR). In my paper, I seek to answer whether the emergence of Large Language Models (LLM) and, in particular, Retrieval-Augmented Generation (RAG) can be considered a turning point in this IR vs. DR debate. Does RAG provide arguments for either side? What should, what can, and what is it practical to expect from a RAG application functioning as a library knowledge system?

Keywords: concept of information, philosophy of information, information retrieval, document retrieval, artificial intelligence, LLM, RAG, meta-information

1. Információ-visszakeresés (IR) és dokumentum-visszakeresés (DR)

Ha könyvtártudományi tanulmányokat folytatva vagy könyvtártudományi kutatásokat végezve azzal a kérdéssel találjuk szembe magunkat, hogy *mikor és milyen körülmények között született meg az információ-visszakeresés terminus*, s könyvtári tudásrendszerhez fordulunk segítségért, kétféle válaszra számíthatunk. Kaphatunk egyfelől közvetlen, azonnali választ, például az alábbi formában: *„Az információ-visszakeresés kifejezést Calvin Mooers amerikai informatikus alkotta meg 1950-ben.”* A könyvtári tudásrendszer által adott másik fajta válasz megbízható, tudományos igényű publikációk listája lehet az információtudomány vagy információtudomány-történet területéről, amelyekben a kifejezés megszületésének körülményei is feldolgozásra vagy legalábbis említésre kerülnek. Egy lehetséges példa ilyen listára:

- Wolfgang G. Stock, Mechtild Stock: *Handbook of Information Science*¹
- Csík Tibor: *Az információtudomány létrejötte és Horváth Tibor információtudományi eszméinek gyökerei*²

1 STOCK – STOCK, 2013:93

2 CSÍK, 2017:59

- Mark Bowles: *The Information Wars: Two Cultures and the Conflict in Information Retrieval, 1945-1999*³
- Ronald R. Kline: *What Is Information Theory a Theory Of?*⁴

Arra, hogy a kétféle válasz közül melyik hasznosabb, melyik elégíti ki jobban a kérdező igényét, általános érvénnyel nem lehet felelni. Bizonyos használati esetekben egyik, másfajta használati esetekben a másik nyújt több segítséget. Az is kétségtelen, hogy elképzelhetőek másfajta válaszok is a fenti két típus valamiféle kombinációjának formájában. A két fenti választípus mibenlétének megértése, különbségeik számbavétele azonban nem hasznontalan és hiábavaló. Ideáltípusokról lévén szó tanulmányozásuk hozzásegíthet a könyvtári tudásrendszerrel szemben támasztható igények, követelmények tisztázásához.

A továbbiakban arra, amit az első esetben végez a könyvtári tudásrendszer, az *információ-visszakeresés* (*Information Retrieval*), arra pedig, amit a második esetben végez, a *dokumentum-visszakeresés* (*Document Retrieval*) kifejezést fogom használni. Ezek az elnevezések természetesen vitathatóak. Az információ-visszakeresés kifejezést sokféle értelemben használják, nem ritkán beleértik ebbe a fogalomba a dokumentum-visszakeresést is.⁵ Mások – talán éppen ezért – az *adat-* vagy *tényvisszakeresés* kifejezést részesítik előnyben az első típus vonatkozásában.⁶ Az elnevezések maguk azonban talán kevésbé lényegesek, ha megfelelően tisztázott az, hogy a kifejezések használója melyik alatt mit ért pontosan. *Információ-visszakeresés* alatt a továbbiakban azt értem, ha a könyvtári tudásrendszer nem pusztán a felhasználó kérdése szempontjából releváns dokumentumokhoz irányítja, hanem közvetlenül és azonnal tényszerű választ ad a megfogalmazott kérdésre. *Dokumentum-visszakeresés* alatt pedig azt, ha a tudásrendszer kitüntetett feladatának a felhasználói kérdés szempontjából legrelevánsabb dokumentumok listájának összeállítását tekintjük.

2. Az ötlet-i vízió és a kritikája

Bár – ahogy ez már a fent kiderült – az *információtudomány* terminus Calvin Mooers nevéhez fűződik, a kifejezés mögött álló gondolat legnagyobb hatású képviselőjének Paul Otlet belga bibliográfust tekinthetjük. Ő volt az, aki az eddig bevett gyakorlattal szembehelyezkedve nem a dokumentumok, hanem „a dokumentumokban található egyedi információk”⁷ számbavételét jelölte meg a könyvtárak legfontosabb feladata gyanánt. 1934-ben írt *Traité de documentation* címet viselő főművében a következőképpen fogalmazza ezt az célkitűzést: „Maguk az adatok jól elkülöníthetők azoktól a dokumentumoktól, amelyek-

3 BOWLES, 1999:158

4 KLINE, 2004:20

5 VAN RIJSBERGEN, 1979:1

6 CAPURRO–HJØRLAND, 2003:384

7 POGÁNYNÉ, 2007:66

ben rögzítették őket. Ezen tények és adatok halmazainak rendszerezett megszervezéséről van szó.”⁸ Az Otlet által *monografikus elv*nek is nevezett vízió szerint egy könyvtári tudásrendszernek nem a létező, nemritkán feleslegesen bőbeszédű és gyakran csupán közismert információkat ismételtető dokumentumok katalogizálásán és osztályozásán kell alapulnia. Az értékes információkat – Otlet szóhasználatával a tényeket és adatokat – elemi formában és a maguk identitásában⁹ ki kell nyerni, ki kell szabadítani a dokumentumokból, és ezekből az önmagában értelmezhető, önmagában teljes információelemekből kell a könyvtári tudásrendszert felépíteni, amely így inkább lesz hasonlatos egy univerzális enciklopédiához, mint könyvtári katalógushoz.

Ehhez a vízióhoz csatlakozott Calvin Mooers a múlt század 50-es éveiben új terminus hozzáadásával. „Az információ-visszakeresés kifejezést alig nyolc évvel ezelőtt volt szerencsém megalkotni egy másik számítástechnikai konferencia alkalmával. A név azóta hosszú utat tett meg. [...] Jelenleg az információ-visszakeresés már többről szól, mint a dokumentumok pusztá megtalálása és rendelkezésre bocsátása; most már az információ felfedezésével és átadásával foglalkozunk, teljesen függetlenül annak dokumentumformájától.”¹⁰ Mooers terminusát a könyvtártudomány szinte azonnal átveszi, mind a mai napig használja.

A mögötte álló otlet-i víziót, azaz könyvtári tudásrendszer IR-rendszerként való felfogásának gondolatát azonban nem fogadta egyöntetű lelkesedés.¹¹ Boyd Rayward, ausztrál kutató például markáns kritikával illeti ezt az elképzelést. Úgy látja, hogy a dokumentumok Otlet szemében csak erősen korlátozott értékkel bírnak, mert ismétlésektől hemzsegek, megfogalmazásuk zavaros, és egyszerre vannak tele igaz és téves állításokkal. „Számára a dokumentumok tartalmának azon aspektusa, amellyel foglalkoznunk kell, a tények. Szinte mindenütt tényekről beszél. Egyértelmű, hogy az általa kifejlesztendőnek vélt új típusú rendszerekkel szemben az elsődleges követelmény az, hogy képesek legyenek felszabadítani az értékes információkat – amit ő tényeknek nevez – az azokat fogva tartó és elrejtő dokumentumokból.”¹² Ez az elképzelés azonban Rayward szerint „tekintélyelvű, redukcionista, pozitivist, leegyszerűsítő – és optimista.”¹³ Arra a nehezen igazolható hitre épül, hogy ami tényszerű és igaz, könnyen azonosítható egy dokumentumban, hiszen ez csupán tartalmának elemzésére és rendszerezésére szolgáló folyamatok szabályozásának és intézményesítésének kérdése.

8 OTLET, 2021:13

9 RAYWARD, 1990:1

10 MOOERS, 1960:229

11 Ld. pl. HJØRLAND, 2026

12 RAYWARD, 1994:247

13 RAYWARD, 1994:247

Birger Hjørland és Rafael Capurro is hasonló álláspontot képvisel, mint Rayward. Szerintük az az elképzelés, amely a könyvtári tudásrendszerek információ-visszakereső rendszeré alakítása mögött áll, idejétmúlt, „meghaladott pozitivizmus.”¹⁴ Robert Fairthorne pedig ironikusan az *információ flogiszton-elméletéről* ír ennek kapcsán, arra utalva ezzel, hogy a dokumentumokban lévő információk visszakeresését propagáló elméletek ugyanúgy tévesen – lényegében egy nem létező dolog feltételezésével – értelmeznek egy jelenséget, ahogy a 17–18. században a flogiszton-elmélet az égést.¹⁵ Ha szöveget olvasunk, nem az történik, hogy valami, ami a dokumentumban van (az információ-flogiszton), „átfolyik” s bekerül a fejünkbe. „Abból, hogy egy rendszer reagál egy ingerre, nem következhetünk valamilyen megmaradó mennyiség létezésére, amely átáramlik rajta. Amikor az ujj meghúzza a ravaszt, nem áramlik semmi az ujjból a golyóba.”¹⁶

Bernd Frohmann Jacques Derrida posztstrukturalista filozófiája mentén fogalmaz meg kritikát Otlet víziójával szemben. Szerinte elemi információk dokumentumokból való kiemelése és kontextusfüggetlen leírása illúzió. A leírás mindig jelekkel (szavakkal) történik, s ezek a jelek nem rendelkeznek azzal a stabilitással, amelyet az otlet-i koncepció feltételez és elvár tőlük. A jelek képlékenyek, mást jelenthetnek helytől és időtől függően. Derrida szerint egy jel éppen attól jel, hogy végtelen különböző kontextusba helyezhető. Így, Frohmann szerint, „derridai probléma merül fel Otlet azon elképzelésével kapcsolatban, miszerint a tények teljes jelenlétük világosságában leírhatóak a tudat számára egy olyan dokumentációs fegyelem révén, amely a tények rögzítését és a jelek közötti kapcsolatok stabilizálását szolgálja.”¹⁷

3. Keresőmotorok, chatbotok és RAG

Ezekben a kritikákban minden bizonnyal sok igazság rejlik, azt állítani azonban, hogy Paul Otlet víziója nem valóban létező, máig jelen lévő felhasználói (kutatói, olvasói) igényre próbálna megoldást találni, túlzás, sőt – megkockáztathatjuk – tévedés volna. Ha másból nem, a nagy nyelvi modellekre épülő generatív mesterségesintelligencia-alapú chatbotok gyors elterjedéséből és rendkívüli népszerűségéből következtethetünk erre. Észre kell ugyanis vennünk, hogy MI chatbotokat (pl. ChatGPT, Gemini) – egyebek mellett – az különbözteti meg a „hagyományos” webes keresőmotoroktól (pl. Google, Bing), hogy míg az utóbbiak dokumentum-visszakeresést végeznek, addig az előbbiek lényegében információ-visszakeresést hajtanak végre. Ha egy webes keresőmotorhoz fordulunk azzal a kérdéssel, hogy *„Milyen hangszerek szólnak Gluck Orfeusz és Euridiké című operájában?”*

14 CAPURRO–HJØRLAND, 2003:385

15 FAIRTHORNE, 1967:711

16 FAIRTHORNE, 1967:711

17 FROHMANN, 2017:85

akkor egy webes keresőmotor online enciklopédiák szócikkeihez, komolyzenei portálok előadás- és lemezkritikáihoz, digitalizált zenetörténeti tankönyvekhez vagy tanulmányokhoz, operaházak webes műsorfüzeteihez és online elérhető operapartitúrákhoz vezető linkek listáját kínálja, azaz olyan dokumentumokhoz vezet el, amelyek relevánsak a kérdés szempontjából, és amelyekből jó eséllyel választ kapok arra, amit tudni szeretnék. Ez pedig nem más, mint amit fent dokumentum-visszakeresésként definiáltunk. Ha viszont ugyanezt a kérdést egy MI chatbotnak teszem fel, akkor azonnal és közvetlenül megkapom azoknak a hangszereknek a listáját, amelyek Gluck a nevezett operában használt. Ez pedig az információ-visszakeresés fenti definíciójának felel meg. A nagy nyelvi modellre épülő MI chatbotok egyre terjedő használatából tehát joggal következtethetünk arra, hogy az információ-visszakeresés létező felhasználói igény.

Fontos azonban azt is látni, hogy a nagy népszerűségnek örvendő chatbotok nem mindenben felelnek meg az ötlet-i vízióknak. A monografikus elvnek ugyanis fontos részét képezi az is, hogy az információ-visszakereséshez használt elemi információk köre miből és hogyan épül fel. Nem pusztán dokumentumokból kiszabadított információknak, hanem *gondosan válogatott dokumentumokból körültekintően kinyert s szakszerűen leírt információknak* kell a visszakeresés alapjául szolgálniuk. Ennek a kritériumnak a népszerű chatbotok nem felelnek meg, hiszen a nagy nyelvi modellek építéséhez kis túlzással válogatás nélkül használtak fel mindenféle szöveges dokumentumot. Hamar ki is derült, hogy naprakész pontosságot és megbízhatóságot igénylő szakértői tudásrendszerként ezek az MI chatbotok nem tudnak működni.¹⁸ Ennek a korlátnak a felismerése hívta életre a RAG-rendszereket, amelyek éppen ebből adódóan az tanulmány első felében tárgyalt vita szempontjából is meghatározó jelentőséggel bírnak.

A RAG (Retrieval Augmented Generation) olyan architektúrát jelent, amely a generatív, nagy nyelvi modellre épülő MI-rendszerbe egy új komponenst (vektoradatbázist) és egy új fázist (retrieval-fázis) illeszt. A RAG-rendszerek építési (indexelési) fázisában ki kell választani azokat a dokumentumokat, amelyeket a rendszer tervezői megbízható információk forrásának tekintenek. Az indexelés során ezeknek a kiválasztott dokumentumoknak a kis darabjait (chunks) numerikus reprezentációkká, ún. vektorokká alakítják, s ezekből a vektorokból speciális vektoradatbázist építenek. Amikor a felhasználó kérdést tesz fel a rendszernek, először a fent említett új (retrieval) fázis indul el, amelynek során rendszer gyors, szemantikus keresést végez az indexelés során felépített vektoradatbázisban. A szemantikus keresés eredményeként a RAG-rendszer megtalálja a megbízhatóként megjelölt, indexelésre kiválasztott dokumentumok kérdés szempontjából legrelevánsabb szövegrészeit. Az LLM-komponens feladata pedig már „csak” annyi, hogy ezekből a szövegrészekből, információelemekből természetes nyelven legenerálja a pontos és tényszerű választ.

18 GAO et al., 2023:1

A RAG-rendszerek tehát már sokkal jobban megközelítik Otlet vízióját, hiszen itt az információ-visszakeresés gondosan és körültekintően kiválasztott dokumentumok szövegrészei alapján történik. A RAG-rendszerek lehetőségeinek és korlátainak számba vétele így az otlet-i koncepció tanulmányozására és értékelésére is alkalmasnak látszik.

4. Információ a dokumentumban és információ a dokumentumról

Mind Otlet, mind a RAG-rendszerek koncepciójának alapmozzanata, hogy az információkat, amelyekre az információ-visszakeresést építeni kívánják, a kiválasztott dokumentumokból nyerik ki. Az információ-visszakeresés tehát mindkét esetben *a dokumentumban lévő információra* épül. Természetesen Otlet tisztában volt azzal, hogy a dokumentumok nem kizárólag információkból állnak, de minden bizonnyal nem tévedünk nagyot, ha azt állítjuk, hogy úgy gondolta, hogy az általa megálmodott információ-visszakeresésnek csak az információkkal van tennivalója. De vajon igaz-e, hogy egy könyvtári tudásrendszer számára csak a dokumentumokban lévő információk relevánsak?

Nézzünk két használati esetet! Az elsőben azért fordul a felhasználó könyvtári tudásrendszerhez, mert gyermekek környezeti neveléséhez a búbos banka hangjáról gyűjt anyagot, a másodikban pedig spanyolországi utazás előtt állva szeretne az Alhambráról tájékozódni. Az első használati eset kapcsán hasonlítsuk össze e két szövegrészletet:

(A₁) „[A búbosbanka] nászhangja öblös, tompa, többször ismételt »up-up-up«.”¹⁹

(B₁) „... ebbe a csendbe lágy tavaszi kiáltással dobol bele a búbos banka.

– Hu-pu-pu, hu-pu-pu...

Ruhája ékes, fején gyönyörű bóbíta, de mást nem tud. – Hu-pu-pu, hu-pu-pu...

Ez a kiáltás mindig közel van, és mindig távol. Egyszerre szól a földön és a levegőben, rászáll a csendre, és csenddé lesz, nincs benne semmi, mégis virágos szomorúság marad utána, mert mindenki tudja, hogyha elhallgat: ledobja szűzi koszorúját a Tavasz, elnehezedik, megbarnul, és tarka, nagy köténye alatt dús anyaságát takargatja a Nyár.”²⁰

19 SVENSSON – GRANT, 2020:467

20 FEKETE, 1971:94

Az első idézet egy terepi madárhathározóból, a második Fekete István *Kele* című regényéből származik. A második használati eset vonatkozásában is vessünk össze két szövegrészletet, az előbbi egy útikönyvből, az utóbbit Száraz Miklós György *Ó, Santo Domingo!* című esszéregényéből idézem²¹.

(A₂) „V. Muhammad remekművét 124 kecses oszlopon nyugvó boltív szegélyezi. A középső kutat 12 zömök kőoroszlán tartja.”²²

(B₂) „Az Alhambra pedig nem építészet, hanem fény. Az ablakok, az áttört díszek, a kagyló- és cseppkőstukkók mind-mind vezetik, terelik, szűrik a fényt, átengedik és megmozgatják, láthatóvá teszik a levegőt. [...] Egyszerre üdítő és lenyűgöző. Csak addig hat, amíg benne vagy. Akkor viszont uralkodik rajtad, nemcsak körülvesz, hanem beléd is költözik.”²³

Talán egyetérthetünk abban, hogy az *A* állításokat joggal tekinthetjük információnak, mind az *A*₁, mind az *A*₂ állítás belefér az információ legszigorúbb meghatározásába is. Megkockáztathatjuk, hogy Otlet monografikus elve szerinti információkinyeréshez is megfelelő alapanyagul szolgálhatnak. Ha egy RAG-rendszer retrieval-fázisa e szövegrészeket keresi elő a vektoradatbázisból, akkor az LLM-fázis megfelelő természetes nyelvű választ fog generálni – amely nem lesz ugyan teljesen azonos a kapott szövegrészlettel, de ennek az eltérésnek (az átfogalmazásnak) a válasz relevanciája és pontossága szempontjából semmi jelentősége nincs.

A *B* állításokról ugyanez viszont nem mondható el. Ha egy RAG-rendszert a búbosbanka hangjáról kérdezve azt a választ kapnánk, hogy „*a búbosbanka hangja egyszerre szól a levegőben és a földön, egyszerre van közel és távol, semmi sincs benne, de szomorúságot okoz*”, akkor joggal lennénk elégedetlenek a rendszer válaszával. S ugyanígy meghökkentőnek és érthetetlennek találnánk, ha egy RAG-rendszer LLM-fázisa a *B*₂ szövegrészlet alapján azt a választ generálná az Alhambrával kapcsolatos kérdésünkre, hogy az „*Alhambra nem építészeti alkotás, hanem fényjelenség, amely beléd költözik, ha benne vagy*”. Ebből arra – az intuíciónknak egyébként tökéletesen megfelelő – következtetésre juthatunk, hogy *B* állítások nem tekinthetőek információnak abban az értelemben, amelyben az *A* állítások azok. Kétségtől van ezeknek is (pozitív) hozadéka a búbosbanka hangjáról, illetve az Alhambráról szóló ismereteink szempontjából, de olvasásuk nem azt a hatást váltja ki

21 Hogy a B₁ és a B₂ szövegrészletek szépirodalmi művekből származnak, nem véletlen. Tanulmányomban arra is fel szeretném hívni a figyelmet, hogy a tudásátadás szempontjából azoknak a könyveknek is jelentős szerepe lehet, amelyekről hajlamosak vagyunk elfeledkezni, amikor (a szűken vett) információközlésről gondolkodunk. Kétségtől van igazsága ugyanis Wittgensteinnek abban, „hogy a költemény, bár az információ nyelvén fogalmazták meg, nem az információközlés nyelvjátékában használatos” [Zettel §160, idézi: FLORIDI, 2011:83], de ehhez azt is hozzá kell tennünk, hogy jóllehet a szépirodalom nem az információközlés nyelvi játékának része, a tudásközvetítésben gyakran játszik fontos szerepet.

22 SÁGHI – DÁVID, 2006:467

23 SZÁRAZ, 2003:162-3, 159

belőlünk, amelyet az *A* állításokhoz hasonló információközlés általában kivált, azt ti., hogy *olvastam valamit, amit eddig nem tudtam, de most már tudom*. Fairthorne nyomán úgy is fogalmazhatnánk, hogy az *A* állítások „információ-flogisztonként” „átfolynak” belénk, a *B* állítások „csak” befolyásolják a gondolkodásunkat, nézőpontunkat, valóságérzékelésünket, változást indukálnak, hozzáárulnak a jelenségek jobb, teljesebb, pontosabb megértésében. Megkülönbözteti a *B* részleteket az *A* állításoktól az is, hogy átfogalmazásukkal elveszik a lényegük, forrásmegjelölés nélkül felhasználói szempontból nehezen értelmezhetőek és értékelhetőek – ami azt is jelenti, hogy a Frohmann által említett derridai probléma ezeknél sokkal határozottabb és egyértelműbben jelentkezik, mint az *A* típusú szövegrészleteknél.

De vajon abból a tényből, hogy a *B* szövegrészletek nem információk, azt a következtetést kell-e levonnunk, hogy érdektelenek egy könyvtári tudásrendszer számára? Semmiképpen sem. Az a tudásszolgáltatás, amelyre a *B* állítások jelentenek példát, legalább olyan fontos, mint az *A* állításokhoz hasonló, szigorúan tényszerű információszolgáltatás. A gyermekek környezeti nevelése szempontjából a *Keléből* vett idézet legalább annyira releváns és hasznos, mint a madárhatározó szikár tényközlése, s bizonyára sokan vannak olyanok is, akik az utazásra való felkészülés tekintetében értékesebbnek találják Száraz Miklós György szubjektív útleírását, mint az útikönyv egzakt információit.

Az *A* típusú szövegrészleteket használva a *dokumentumban lévő információból* megfelelő választ tud generálni egy információ-visszakereső rendszer. *B* típusú szövegrészletek kapcsán azonban nem a *dokumentumban lévő információra*, hanem a *dokumentumról szóló információkra*, azaz *metainformációra* építve kell a választ kigenerálnia²⁴:

(M₁) Fekete István *Kele* című művében lírai leírást találunk a búbosbanka hangjáról.

(M₂) Száraz Miklós György *Ó, Santo Domingo!* című esszéregényében az Alhambra Oroszlános udvaráról is olvashatunk szubjektív útleírást.

Az ilyen válaszok azonban már nem egy információ-visszakereső rendszer válaszai, hanem egy dokumentum-visszakereső rendszeré, hiszen releváns dokumentumokhoz irányítanak, s nem közvetlen, tényszerű választ közölnek a feltett kérdésre. Ami más-képpen fogalmazva azt jelenti, hogy annak, hogy *A* típusú dokumentumokból információ-visszakereső rendszert építsünk, van létjogosultsága, *B* típusú dokumentumokból azonban csak dokumentum-visszakereső rendszert lehet és szabad építeni. Egy könyvtári tudásrendszernek – legyen az akár hagyományos, akár MI- vagy RAG-rendszer – különbséget kell tudni tenni *A* és *B* típusú dokumentumok, dokumentumrészletek között, és ennek függvényében kell információ- vagy dokumentum-visszakereső rendszerként működni.

Minden bizonnyal kiválóan lehet amellet érvelni, hogy az *A* típusú információszolgáltatás előnyére válhat egy könyvtári tudásrendszernek, de joggal gondolhatja azt is bárki, hogy

24 idézi: HJØRLAND, 2026:4

az adott szakterület (ornitológia, turizmus) domain-specifikus tudásrendszere legalább annyira alkalmas eszköz lenne erre. Az azonban, hogy az *M* típusú metainformáció-szolgáltatás számára aligha találhatunk megfelelőbb, autentikus helyet egy könyvtári tudásrendszernél, azt gondolom, nehezen vitatható.

5. Konklúzió

A RAG-rendszerek több fontos tekintetben rokonságot mutatnak az ötlet–mooersi-i vízió információ-visszakereső rendszereivel, hiszen

- gondosan kiválasztott dokumentumokból dolgoznak
- a *dokumentumokban lévő információra* alapozva, s
- közvetlen, tényszerű választ adnak
- a dokumentumok szövegének átfogalmazásával.

Nem véletlen, hogy a könyvtári informatika egyik legmeghatározóbb trendje napjainkban a RAG-technológia integrálása, s hogy a RAG-technológiára épülő könyvtári tudásrendszer kérdése a kutatási és fejlesztési törekvések fókuszába áll.²⁵ A RAG-rendszerek működésének – akár csak gondolat kísérlet révén történő tanulmányozása segítséget jelenthet a könyvtári információ-visszakeresés lehetőségeinek, korlátainak jobb megértéséhez is.

Felhasznált források

- Bevara, R. V. K. – Lund, B. D. – Mannuru, N. R. – Karedla, S. P. – Mohammed, Y. – Kolapudi, S. T. – Mannuru, A. (2025). Prospects of Retrieval Augmented Generation (RAG) for Academic Library Search and Retrieval. *Information Technology and Libraries*, 44(2). <https://doi.org/10.5860/ital.v44i2.17361>
- Bowles, M. (1999). The Information Wars: Two Cultures and the Conflict in Information Retrieval, 1945-1999. In: Bowden, M. E. – Hahn, T.B. – Williams, R.V. (eds.). *History and Heritage of Science Information Systems*. Medford: Information Today, 155-166.
- Capurro, R. – Hjørland, B. (2003). The Concept of Information. *Annual Review of Information Science and Technology*. 37(1), 343–411. <https://doi.org/10.1002/aris.1440370109>
- Csik T. (2017). Az információtudomány létrejötte és Horváth Tibor információtudományi eszméinek gyökerei. *Könyvtári Figyelő*. 27(1), 56–67.
- Fairthorne, R. A. (1967). Morphology of „Information Flow”. *Journal of the Association for Computing Machinery*. 14(4), 710–719. <https://doi.org/10.1145/321420.321430>
- Fekete I. (1971). *Kele*. Budapest: Móra.
- Floridi, L. (2011). *The Philosophy of Information*. Oxford, Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199232383.001.0001>

25 Ld. pl. BEVARA–LUND–MANNURU–KAREDLA–MOHAMMED–KOLAPUDI–MANNURU, 2025

- Frohmann, B. (2017). The Role of Facts in Paul Otlet's Modernist Project of Documentation. In: Rayward, W. Boyd (ed). *European Modernism and the Information Society*. Aldershot: Routledge, 75-88. <https://doi.org/10.4324/9781315580951-5>
- Gao, Y. – Xiong, Y. – Gao, X. – Jia, K. – Pan, J. – Bi, Y. – Dai, Y. – Sun, J. – Guo, Q. – Wang, M. – Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey. *ArXiv*, <https://doi.org/abs/2312.10997>
- Hjørland, B. (2026). Information Retrieval or Document Retrieval? *Journal of Association for Information Science and Technology*. 7(5), 714-726. <https://doi.org/10.1002/asi.70057>
- Kline, R. (2004). What Is Information Theory a Theory of? In: Rayward, W. B. – Bowden, M. E. (eds). *The History and Heritage of Scientific and Technical Information Systems*. Medford: Information Today, 15–28.
- Mooers, C. N. (1960). The Next Twenty Years in Information Retrieval. *American Documentation*. 11(3), 229–236. <https://doi.org/10.1002/asi.5090110306>
- Otlet, P. (2021). *Traité de documentation*. La Plaine-Saint-Denis: Éditions des maisons des sciences de l'homme associées. <https://doi.org/10.4000/books.emsha.482>.
- Pogányné R. G. (2007). Törekvések a világ tudásának egyetemes számbavételére Paul Otlet és Henri La Fontaine munkásságában. *Könyvtári Figyelő*. 53(1), 55–73.
- Rayward, W. B. (1990). Introduction. In: Otlet, Paul (1990), *International Organisation and Dissemination of Knowledge: Selected Essays of Paul Otlet*. Amsterdam and New York: Elsevier, 1–10.
- Rayward, W. B. (1994). Visions of Xanadu: Paul Otlet (1868–1944) and Hypertext. *JASIST*. 45(4), 235–250. [https://doi.org/10.1002/\(SICI\)1097-4571\(199405\)45:4<235::AID-ASL2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199405)45:4<235::AID-ASL2>3.0.CO;2-Y)
- Sághi B. – Dávid K. (2006). *Spanyolország*. Budapest: Panemex-Grafó.
- Stock, W. G. – Stock, M. (2013). *Handbook of Information Science*. Berlin: De Gruyter Saur. <https://doi.org/10.1515/9783110235005>
- Svensson, L. – Grant, P. J. (2005). Madárhatározó. *Európa és Magyarország legátfogóbb terepi határozója*. Budapest: Park.
- Száraz M. Gy. (2003). *!Ó, Santo Domingo!* Budapest: Magyar Könyvklub.
- van Rijsbergen, C. J. (1979). *Information Retrieval*. London: Butterworths. Forrás: <http://www.dcs.gla.ac.uk/MKeith/Preface.html>
- Zins, Ch. (2007). Conceptual Approaches for Defining Data, Information, and Knowledge: Research Articles. *Journal of the American Society for Information Science and Technology*. 58(4), 479–493. <https://doi.org/10.1002/asi.20508>