

Magyar nyelvű internetes keresők vizsgálata a relatív teljesség és a pontosság szempontjából

TÓTH Erzsébet

Előjáróban

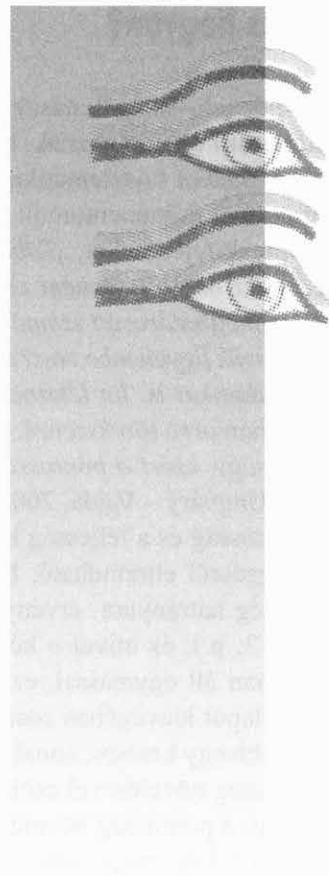
A html-dokumentumok gyors elérését a szerveroldali keresőszolgáltatások biztosítják. E szolgáltatások lehetnek globálisak, ha gyűjtőkörük a web egészére terjed ki, lehetnek nemzetiek, ill. regionálisak, vagy speciális tartalmú, esetleg speciális típusú html-dokumentumot szolgáltatók. Az információkereső rendszerek hatékonysága több szempont szerint vizsgálható, így például a keresési eszközök gazdagsága, az átvizsgált adatbázisok nagysága, a keresési eredmények relevanciája stb. szerint.

Egy keresőszolgáltatás hatékonyságának jellemzésére jómagam a keresés pontosságát és

a keresés teljességét kifejező mutatókat tartom leginkább jellemzőnek, ezért e két szempont megvalósulását elemeztem öt magyar nyelvű keresőszolgáltatásban.¹

Pontosság alatt a releváns dokumentumoknak a visszakeresett dokumentumok összességén belüli arányát, *teljesség* alatt pedig a megtalált releváns dokumentumoknak az összes (akár talált, akár nem) releváns dokumentumhoz viszonyított arányát értettem (Ungváry, 2001.196–197. p.).

1. Jelen írás a szerző Az internetes keresők lekérdezési hatékonyságának vizsgálata című PhD-dolgozata részfejezetének lapunk számára átdolgozott, rövidített vátozata.



Mit és hogyan?

Az indexelő-szolgáltatások (keresőgépek, katalógusok) úgy dolgoznak, hogy meghatározott időközönként szisztematikusan átnézik a hálózaton lévő dokumentumok jellemző adatait és az indexkifejezéseket. „Többségük nem a teljes szöveget nézi át, hanem csak a dokumentum elején meghatározott számú (pl. 20–30) sort, és egy részük figyelembe veszi a címfejlébe szereplő metaadatokat is, ha léteznek. A szolgáltatások általában arra törekszenek, hogy a teljesség legyen nagy, ezért a pontosságról eleve nincsen szó” (Ungváry – Vajda, 2002. 111. p.).

A pontosság és a teljesség között fennálló összefüggésről elmondható, hogy a pontosság a teljesség hátrányára érvényesül (Brin – Page, 1998. 3. p.), és mivel e két fogalom fordított arányban áll egymással, ezért az ideális keresési állapot lényegében sosem érhető el. Minél teljesebb egy keresés, annál pontatlanabb, mert a teljesség növelésével csökken a pontosság és viszont: a pontosság növelése a teljesség csökkenését vonja maga után (Horváth – Sütheő, 2003. 180. p.).

Elemzéseim a következő öt magyar nyelvű keresőszolgáltatásra irányultak: az AltaVizslára, a Heurékára, az Origo-Vizslára, a Kurzorra és a Góliátra. (Nem vizsgáltam viszont a több nyelvű, ill. a magyar nyelven is használható egyetemes keresőszolgáltatásokat.)

A pontosságot a használat és a hatékonyság összefüggéseiben elemeztem négy keresőkérdés feltételével és az első hatvan találatban. Azért csak az első hatvan találatig, mert a Google fejlesztőinek adatai meggyőztek arról, hogy a használók csak az első tíz-húsz találatot veszik figyelembe. Nem véletlenül törekednek a fejlesztők arra, hogy a keresőkérdésekre a leginkább releváns találatok a találati lista legelején jelenjenek meg.

A használat szerinti elemzés megmutatja, hogy a vizsgált keresők hány releváns találatot tudnak visszakeresni az első, a második és a harmadik húszas találati blokkban. Ez fontos jelzés a hasz-

nálók számára, mert ennek alapján eldönthetik, melyik keresőt érdemes használniuk a releváns információk felkutatására.

A találatok relevanciáját a közölt találati lista sorrendje alapján állapítottam meg. A találati pontosságra kiszámítottam egy százaléktérteket, ami megmutatta, hogy hány százalékban kaptam releváns találatokat az adott keresőkérdésekre a vizsgált első, második és harmadik húsz találatból. Megnéztem, hogy ezekből milyen következtetést lehet levonni a releváns találatok számára a találati halmaz hátralévő részére.

A teljesség vizsgálatát a szakirodalomból ismert relatív teljesség alapján végeztem el, ami a következőképpen számolható ki: az adott kereső által visszakeresett releváns dokumentumok száma osztva az összes vizsgált keresők valamelyike által megtalált releváns dokumentumok számával (Clarke – Willett, 1997. 185–187. p.). Lefuttatott keresésem a teljesség mérésére a csiperkegombával foglalkozó forrásokra irányult. A „csiperkegomba” témakörre irányuló kereséskor a releváns találatok halmaza huszonegy releváns dokumentumból állt. A 21-es halmazt a vizsgált keresők összes visszakeresett találatából nyertem. Ezután nyomon követtem a témára vonatkozó hiperhivatkozásokat, valamint a kapcsolódó szinonimákat is. A kiválasztott releváns weboldalak referenciaként kerültek bele a halmazba és bárholonnan felkutathatók voltak a weben. Ezt követően megnéztem, hogy az egyes keresők mennyire voltak képesek megtalálni a halmaz releváns dokumentumait a saját találati halmazukban. Ezek alapján kiszámoltam egy százaléktérteket a relatív teljességre vonatkozóan, ami kifejezte, hogy az egyes keresők a vizsgált halmaz releváns dokumentumainak hány százalékát találták meg.

Egy weboldalt akkor minősítettem relevánsnak, ha tartalmazta a megadott keresőkérdést és/vagy annak más egyéb szinonimáit; ha technikailag bármikor elérhető volt; ha magyar nyelven jelent meg; és ha tartalmában és teljes terjedelmében szorosan kapcsolódott a megjelölt témához.

Az utóbbi feltétel teljesülése érdekében a weboldalak tartalmát minden esetben külön is

ellenőriztem, és formai szempontok szerint is meggyőződtem megbízhatóságukról és hitelességükről. Ha a definiált feltételek együttesen nem álltak fenn, akkor a weboldal vizsgálatom számára nem minősülhetett relevánsnak.

A keresőkérdések szinonimáival kapcsolatban megjegyzem, hogy a szemantikus keresők elméletileg megtalálják egy téma szűkebb vagy tágabb értelemben vett szinonimáit, de az eldöntendő kérdés lett számomra, hogy mikor tekinthetők egy szinonimát a téma szempontjából is relevánsnak. A relevancia értelmezése nagy körültekintést igényelt.

A vizsgálatban szereplő keresőeszközök

Vizsgálatomhoz azért választottam a már említett öt magyar nyelvű keresőt (Heuréka, Origo-Vizsla, Kurzor, Góliát és az AltaVizsla), mert azon túl, hogy ezek a leggyakrabban használt keresőgépek a magyar használók körében, elmondható, hogy találatsszolgáltatásuk megbízható, rövid válaszdőn belül képesek megfelelő mennyiségű találatot adni a lefuttatott keresésekre. Kritikaként annyit jegyeznek meg, hogy nem szűrik ki az ismétlődő találatokat a találatthalmazokban, ezért a felbukkanó duplum oldalak növelik a zajos találatok számát, (ezeket irreleváns dokumentumoknak tekintettem).

Keresőkulcs vagy *keresőkulcs érték* alatt azokat a használók által megadott releváns kulcsszavakat értem, amelyekre a keresőgépek a weboldalak visszakeresését végezték. Használok még a *témakör kifejezést* is. Egy adott témakört egy speciális keresőkulcs érték jelölt (például a „szökőár” keresőkulcs érték fejezi ki a „szökőár” témakört). A továbbiakban a témaköröket, a keresőkulcsokat, illetve a keresőkulcs értékeket relációs jelek között dőlt betűvel jelölöm.

A vizsgálat során használt keresőkulcsok

Vizsgálatomat 2006 januárja és júniusa között végeztem el manuális úton, hogy kiküszöböljem az automatizálásból eredő téves értelmezéseket és hibákat, összesen öt témakörben. Választott témáim a következők voltak: <csiperkegomba>, <Rákóczi-szabadságharc>, <evészavar>, <kegyhely>, <szökőár>.

Az öt téma közül csak egy irányult a teljesség vizsgálatára (<csiperkegomba>), a fennmaradó négy téma (<Rákóczi-szabadságharc>, <evészavar>, <kegyhely>, <szökőár>) a pontosság tanulmányozására szolgált.

A teljesség vizsgálatára olyan összetett keresőkérdést fogalmaztam meg, amely teljes mértékben figyelembe vette a megadott témakör sajátosságait, míg a pontosság vizsgálatokor először rögzítettem az egyes keresőkérdésekre kapott összes találat számát, melyre azért volt szükség, mert az összes találati szám alapesetben befolyásolja a releváns találatok számát a találatthalmaz további (hátralévő) részében.

A lekérdezett dokumentumokat teljes weboldalként saját gépemre mentettem az Explorer böngészőn keresztül. Tekintve, hogy a pontosságot csak az első hatvan találatig tanulmányoztam, ezért értelemszerűen csak ezt a találatmennyiséget mentettem el. A mentéseket logikus felépítésű könyvtár struktúrába rendeztem, melyben a könyvtárak elnevezései a vizsgált keresőknek, az alkönyvtárak nevei pedig a keresőkérdéseknek feleltek meg. A találatok tartalmát a kérdések lefuttatása után egyenként is átnéztem, mert a relevancia megállapítása megkövetelte az alapos ellenőrzést. A relevancia megállapítására nem hoztam létre külön kategóriákat, de a találatok teljes szövegéből kiemeltem azt a részt, aminek alapján az adott találatot relevánsnak minősítettem. Az 1. függelékben közlöm az általam relevánsnak tekintett találatokat <csiperkegomba> témakörben, ahol feltüntetem, hogy melyik kereső kereste vissza az adott találatot a vizsgálat során. Továbbá jelölöm, hogy melyek azok a re-

leváns, bárholnan felkutatott hiperhivatkozások, amelyeket egyik vizsgált kereső sem talált meg. A 2. függelékben összefoglalom az egyes keresőkulcsokra lekérdezett releváns és összes találatok számát.

A keresőkulcsok kiválasztása

A vizsgálat megtervezésekor arra törekedtem, hogy megtaláljam azokat a keresőkifejezéseket, amelyek kielégítik a vizsgálat célkitűzéseit. Először próbakereséseket futtattam le, majd az ekkor szerzett tapasztalatokra támaszkodva a későbbiekben pontosítottam szempontjaimat. A próbakeresések egyértelművé tették, hogy olyan keresőkérdéseket célszerű használni a vizsgálatához, amelyek minél pontosan behatárolják a témát. Például a <szökőár> keresőkulcs sok használható találatot hozott. Egy túl általános kifejezés fölöslegesen sok irreleváns találatot eredményez. (Ezért hagytam ki pl. a <rajzfilm> keresőkulcsot, mert túl sok zajos találatot kaptam.)

Ezután próbakereséseket végeztem többszavas keresőkérdésekre is keresési operátorok („...” jel, ÉS, + jel) segítségével. Kiválasztottam a <táplálkozási zavar> keresőkérdést, amely ugyanabban az értelemben használatos a pszichológiában, mint az <evészavar>. Erre a kifejezésre keresve túl sok irreleváns találatot kaptam (pl. sok olyan oldalt szolgáltatnak a keresők, amelyek általában az egészséges életmóddal és az étrenddel foglalkoztak). Ezért elvettem a <táplálkozási zavar> többszavas keresőkulcs használatát és helyette az <evészavar> keresőkérdést alkalmaztam, amellyel lényegesen több használható találatot kaptam. Kipróbáltam még a <testkép zavar> keresőkulcsot is, amely ugyanabban az értelemben használatos kifejezés a pszichológiában, mint az előző két keresőkérdés. Az így feltett keresés témába vágó találatokat eredményezett ugyan, de a találatok kis száma miatt mégsem vettem figyelembe ezeket. Kis találatszámúnak a hatvan találat alatti halmazokat tekintettem. Mivel e találatmennyiségnek a megléte a pontosság tanul-

mányozásának alapkritériuma volt, ezért ehhez ragaszkodnom kellett a vizsgálat során.

A másik kifejezés: a <Rákóczi-szabadságharc> keresőkulcs érték egyértelműen meghatározott egy magyar történelmi eseményt, ezért nem volt meglepő, hogy sok releváns dokumentumot kaptam. Összetett alakja ellenére egzakt keresőkérdésnek minősült.

A pontosság vizsgálatokor alapkövetelménynek tekintettem, hogy csak olyan témákat vegyek figyelembe, amelyekre elegendő számú releváns találatot kapok. Ezért pl. az <organikus építészet> keresőkulcs nem felelt meg a célkitűzéseimnek, mert kevés releváns weboldalt eredményezett. Ezzel szemben a <kegyhely> keresőkulcs értékre gazdag és releváns dokumentumanyagot kaptam, ezért ezt megfelelő keresőkérdésnek tekintettem.

A teljesség vizsgálatokor nehezebb dolgom volt, mert utána kellett nézmem az adott téma specifikus jellemzőinek. A <csiperkegomba> témakörhöz kapcsolódóan a következő keresőkulcs értékeket határoztam meg: <sampinyon>, <termesztett csiperke>, <kétspórás csiperke>, <agaricus bisporus>. Igyekeztem olyan keresőkérdéseket választani, amelyek pontosan körülírták a fogalmat. A keresőkulcsokból végül – egy kivételével – kizártam a csiperkegomba különböző fajait és latin nyelvű megnevezéseit. Elektronikus növény- és állatrendszertani tezauszban, valamint egy gombákkal foglalkozó szakkönyvben jártam utána, milyen egyéb szinonimákat szoktak használni a szakértők a csiperkegombára. (Albert–Locsmándi–Vasas, 1995. 20. p.; Taxaurusz, 2000). Ekkor találtam rá a „sampinyon” szinonimára, de észrevettem, hogy ezt a megnevezést csak a kétspórás csiperkére és a természetett csiperkére használják. (Megjegyzem, hogy a csiperkegomba legismertebb természetett fájának a kétspórás csiperkét tekintik általában.) Ezért a <sampinyon>, a <kétspórás csiperke> és a <termesztett csiperke> keresőkulcsokat is releváns keresőkérdésként kezeltem. Megtaláltam a kétspórás csiperkének az <agaricus bisporus> latin nyelvű megnevezését is, amit szintén releváns keresőkulcsnak értékel-

tem. Mindegyik keresőkulcsot külön-külön lefuttattam az egyes keresőkön. A kapott találatokból – tartalmi ellenőrzés után – kiválasztottam azokat a relevánsakat, amelyeket még egyszer külön is ellenőriztem a keresőkben.

A pontosság vizsgálata

A „Rákóczi-szabadságharc” témakör

Ebben a témában túl sok releváns forrást találtam az interneten, melynek az a magyarázata, hogy akkoriban (2006) emlékeztünk meg a szabadságharc elindulásának 300. évfordulójáról. Az évfordulóval foglalkozó weboldalak közül azokat tekintettem relevánsnak, amelyekben részletes leírást is találtam a szabadságharc eseményeiről, illetve vezéralakjairól. Relevánsnak vettem a szabadságharc különböző aspektusait bemutató tudományos igényességgel készült szakcikkeket, tanulmányokat, konferencia-előadásokat. Voltak relevánsnak minősíthető érettségi tételek is a témáról, és találtam olyan írásokat is, amelyek a szabadságharc egy adott településhez kapcsolódó eseményeit mutatták be. Ezeket az oldalakat a helyismereti kutatásra tekintettel minősítettem relevánsnak. Nem vettem fel viszont olyan weboldalakat, ahol csak néhány mondatban hivatkoztak a Rákóczi-szabadságharcra mint a település életében lezajlott eseményre. Relevánsnak ítéltam meg a Magyar Elektronikus Könyvtárban található digitalizált kézikönyvek vonatkozó fejezeteit, valamint az elektronikus lexikonok (például Magyar Virtuális Enciklopédia, Wikipédia) szócikkeit is.

A „szökőár” témakör

Ezzel a témával kapcsolatban ugyancsak sok hírt, cikket, ismeretterjesztő szintű írást találtam. Nem minősítettem relevánsnak azokat az oldalakat, amelyek egyes országok és neves cégek jótékonsági célú adományozásairól számoltak be a katasztrófától sújtott területeken. Relevánsnak vettem viszont azokat az oldalakat, amelyek

a katasztrófák következményeiről, valamint a nemzetközi humanitárius szervezetek különböző segélyezési akcióiról számoltak be. Relevánsnak fogadtam el azokat az oldalakat is, amelyek ismeretterjesztő szinten tárgyalták pl. egy szökőár jellegzetességeit, kialakulásának okait, előrejelzésének lehetőségeit. Itt is mellőztem a témáról szóló rövid terjedelmű híryanagokat.

Az „evészavar” témakör

Az <evészavarok> fogalma alatt elsősorban az „anorexia nervosa”-t és „bulemia nervosa”-t szokták érteni. Mindkét típusú táplálkozási probléma leggyakrabban fiatalkorú lányoknál fordul elő. A téma szempontjából relevánsnak számítottak azok a dokumentumok, amelyek a két betegség tüneteit együttesen vagy külön-külön ismertették. Ebből adódóan egyértelműen megállapítható volt, hogy a dokumentum a megadott témáról szól-e vagy sem. Számos orvostudományi szakfolyóiratban megjelent cikket, továbbá néhány ismeretterjesztő szintű weboldalt találtam a témáról, amelyek bemutatták a betegség tüneteit, és leírták terápiás kezelését. A források másik köre azokra a társadalmi, kulturális, biológiai és pszichológiai tényezőkre hívta fel a figyelmet, amelyek az evészavarok kialakulásához vezetnek.

A „kegyhely” témakör

Ezzel a témával kapcsolatban ugyancsak sok, többnyire ismeretterjesztő dokumentumot találtam az interneten. Több központi gyűjtőoldalt találtam, de ezeket általában nem tekintettem relevánsnak, csak akkor ha egy adott kegyhely történetét, illetve az ott meghonosodott búcsújárás történetét tekintették át. A keresők találtak olyan oldalakat is, amelyek a „búcsújáróhely” szinonimát tartalmazták a kegyhely keresőkérdésre. Nem fogadtam el relevánsnak a híreket és a témát csak érintőlegesen tárgyaló helyeket.

A teljesség vizsgálata

A teljesség vizsgálatokor először a releváns találatok halmazát kellett létrehoznom a lekeresett összes találatból a már említett <csiperkegomba> témakörben. Ezután megnéztem, hogy az egyes keresők milyen arányban tartalmazták a releváns halmaz dokumentumait.

„Csiperkegomba” témakör

A halmazba kerülő releváns dokumentumok kiválasztásakor releváns dokumentumnak tekintetem minden olyan tudományos kutatásról szóló szakmai beszámolót, amely valamilyen speciális szempontból vizsgálta a csiperkegomba természetét. Doktori tézisek, disszertációk, szakdolgozatok, valamint szakmai jellegű pályázati anyagok reprezentálták ezt a kört. Nem tettem különbséget a különböző tudományterületek megközelítései között, mindegyiket egyenrangúnak tekintetem. Relevánsnak vettem továbbá minden olyan írást, amely ismeretterjesztő szinten mutatta be a csiperkegomba-termesztést, vagy tartalmazta a csiperkegombának mint növényfajnak az ismertetőjegyeit. Figyelmen kívül hagytam viszont a csiperkegomba egyes fajaira vonatkozó weboldalakat, mert azokra külön nem kerestem a keresőkérdésekben. Kivételt képezett a kétspórás csiperke gombafaj, valamint latin nyelvű megnevezése, amelyekre indokoltnak tartottam a keresést. Nem tekintetem relevánsnak azokat a weboldalakat, amelyek csak 1-2 mondatban tárgyalták a témát. Ha egy dokumentumban több gombafaj szerepelt és csak röviden utalt a csiperkegombára, akkor azt is kizártam, (pl. a gombaszagokról vagy a gombamérgezésről szóló dokumentumokat). Ha a dokumentum átfogó képet nyújtott a gombatermesztés helyzetéről a magyar mezőgazdaságban, azt sem tekintetem relevánsnak, mert nem csak a csiperkegomba termesztésével foglalkozott. Kihagytam még az angol nyelvű weboldalakat és a nem létező oldalra történő hivatkozásokat (Berners-Lee – Fielding – Frystyk, 1996, RFC1945) is.

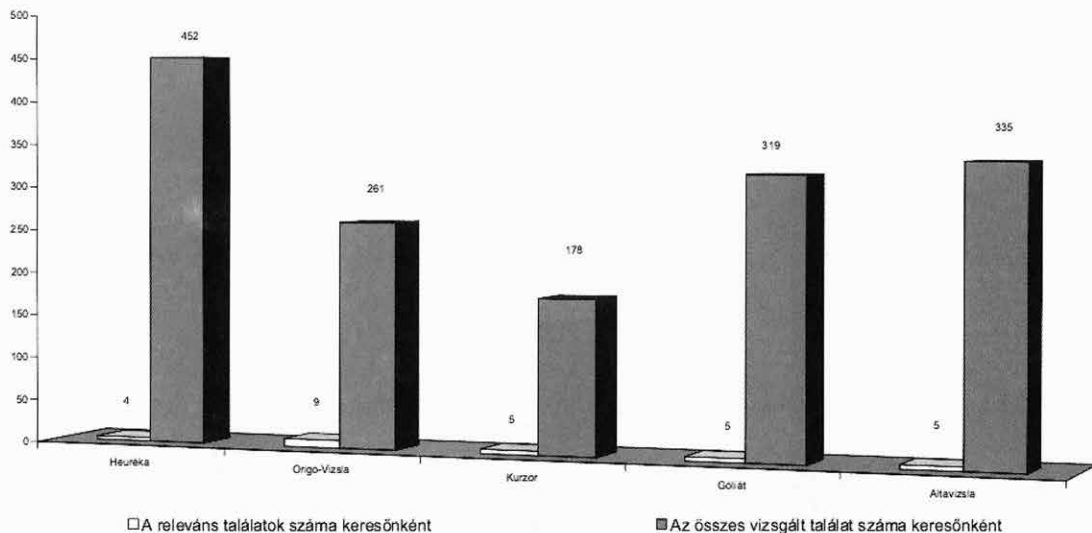
Amikor azt elemeztem, hogy az egyes keresők hány találatot adtak vissza a releváns dokumentumok halmazából, a következő esetekkel találkoztam: ha a halmazban lévő weboldal ugyanazon az URL-en volt lekérdezhető, mint ahol azt korábban a kereső felkutatta, akkor ezt megtalált dokumentumnak minősítettem (természetesen ez az eset fordult elő a leggyakrabban); ha az oldal párhuzamosan több URL-en volt elérhető, a kereső pedig ezek közül csak az egyik helyen találta meg a dokumentumot, akkor azt is felkutattott releváns találatnak vettem; ha a kereső a dokumentumnak egy olyan korábbi URL-jét indexelte, ahol már ténylegesen nem volt elérhető és csupán a „cache”-ben tárolt változatát szolgáltatta, akkor azt kizártam (mert a releváns dokumentumok halmazát létező, elérhető dokumentumokra mutató halmaznak vettem).

A teljességre kapott eredmények

A <csiperkegomba> témakör esetében rögzítettem, hogy összesen hány találatot adtak az egyes keresők a megadott keresőkulcsokra. A keresőkérdésekre kapott összes találatszámot összeadtam, ezáltal jó gyakorlati mérőszámot kaptam a keresők által szolgáltatott összes találat számára. A közösen előforduló duplum weboldalakat számát az egyes keresőkulcsokra nagyjából azonosnak vettem. Ezt követően az öt kereső összes találatainak a számát és az általam ténylegesen relevánsnak minősített találatok számát elemeztem (ld. 1. diagram).

A releváns találatok száma és az összes találatok száma

Az 1-es diagramon megfigyelhető, hogy az összes találatzámnak milyen elhanyagolhatóan kis töredéke minősült relevánsnak a téma szempontjából. Megfigyelhető még egy összefüggés is: egy nagyobb, az összes találatból álló halmaz nem feltétlenül eredményez több releváns dokumentumot a kereső számára. Ez a megállapítás a Heuréka keresőre vonatkozik, mert a többi keresőszolgáltatás több releváns találatot nyújtott.



1. diagram

A releváns találatok száma és az összes találatok száma a „csiperkegomba” keresőkérdésre

A következő lépésben megvizsgáltam, hogy az egyes keresők pontosan hány dokumentumot kerestek vissza az általam létrehozott releváns találatok halmazából, ami huszonegy weboldalt tartalmazott (1. táblázat).

A relatív teljesség

Keresők	Megtalált releváns találatok
Heuréka	4
Origo-Vizsla	9
Kurzor	5
Góliát	5
Altavizsla	5
Átlaga:	5,6

1. táblázat

A keresők által visszakeresett releváns találatok száma és azok kiszámított átlagértéke

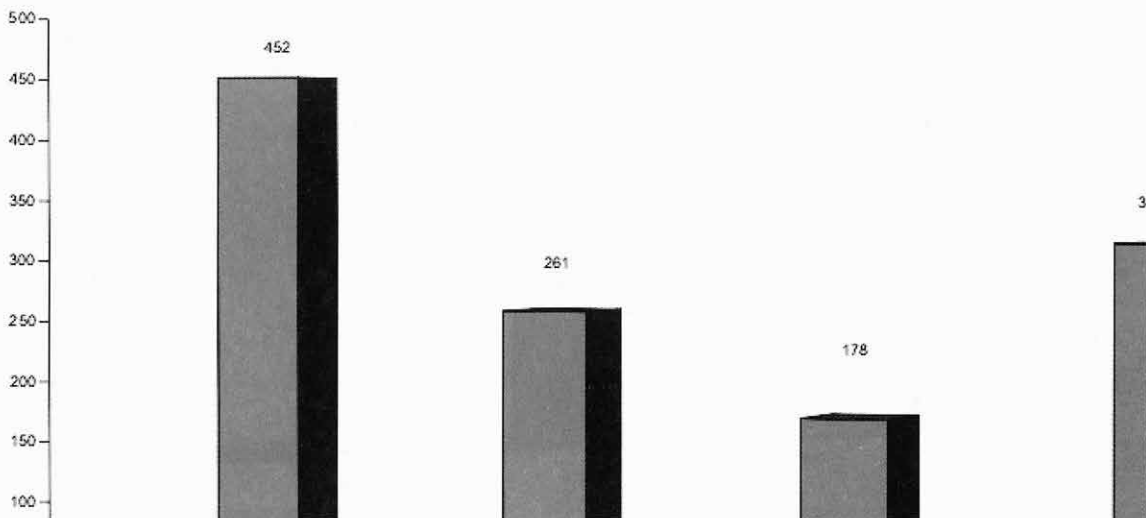
A táblázat azt mutatja, hogy a legtöbb dokumentumot az Origo-Vizsla kereső találta meg a releváns találatok halmazából. A többi kereső közel azonos módon teljesített. Tehát a keresők átlag-

osan 5,6 releváns dokumentumot kutattak fel a weben a <csiperkegomba> témában, ami a vizsgált halmazban lévő huszonegy találat 26,66%-ának felelt meg. Megjegyzem, hogy a bárhol felkutatott releváns hiperhivatkozásokat – melyeknek száma összesen 8 volt a vizsgált halmazban – egyik magyar nyelvű kereső sem kereste vissza. Ezután kiszámoltam egy százalékként a relatív teljességre, ami megmutatta, hogy az egyes keresők a releváns találathalmazból a releváns dokumentumok hány százalékát keresték vissza. A 2. diagram ezt az értéket mutatja be az egyes keresőkre lebontva.

A 2-es diagramon látható, hogy az Origo-Vizsla kereső volt a legeredményesebb: a vizsgált találathalmaz releváns dokumentumainak 42,85%-át kutatta fel a weben. A rangsorban ezután a Kurzor, a Góliát és az Altavizsla következtek 23,80%-kal, az utolsó helyre a Heuréka került 19,04%-kal.

A pontosságra kapott eredmények

A pontosság vizsgálatához a négy különböző keresőkulcsot futtattam le, és a kapott találati listákban az első hatvan találat relevanciáját



2. diagram

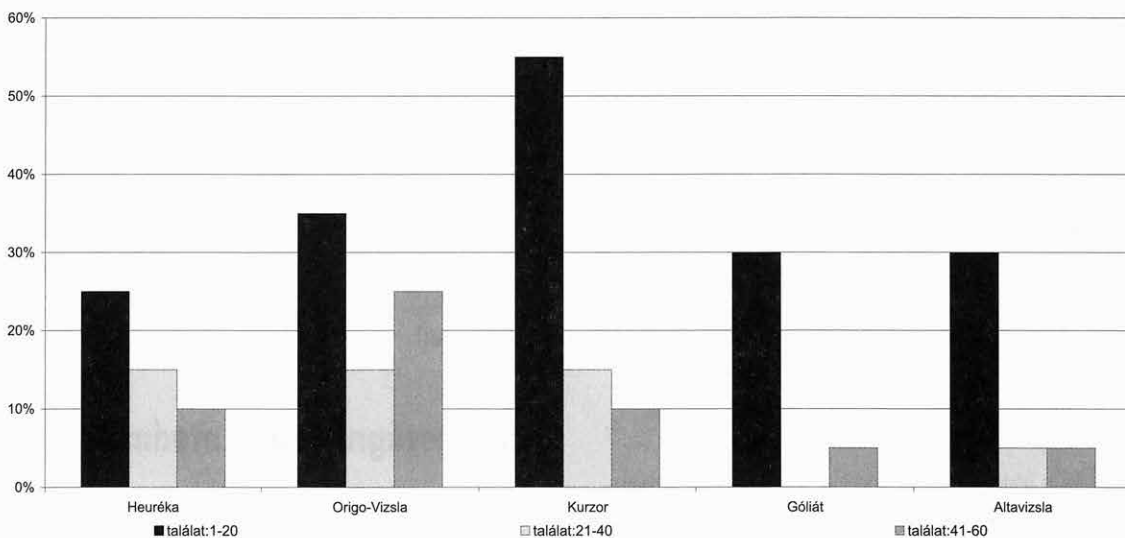
A keresők relatív teljesség szerinti százalékos megoszlása

vizsgáltam, majd kiszámítottam egy százalékértéket, ami azt mutatta, hogy hány releváns találatot kaptam az első, a második és a harmadik hús visszakeresett találatból.

A továbbiakban bemutatom, hogyan alakultak az egyes keresőkulcsokra kalkulált találati pontosság értékei keresőnként az első, a második és a harmadik húsas találati blokkban.

A találati pontosság változása keresőkulcsenként

A 3-as diagram megmutatja, hogy a <Rákóczi-szabadságharc> keresőkulcsra mind az öt kereső az első hús találatban szolgáltatva a legtöbb témába vágó találatot. A Kurzornál volt a legnagyobb a releváns találatok száma, a második az Origo-Vizsla lett. A Góliát és az Altavizsla azonos mennyiségű releváns találatot adott az első húsban, míg a legkevesebb releváns találatot a Heuréka produkálta. A Heuréka és a Kurzor keresők szigorúan monoton csökkenő tendenciával hozták a releváns talátaikat az első hatvan



3. diagram

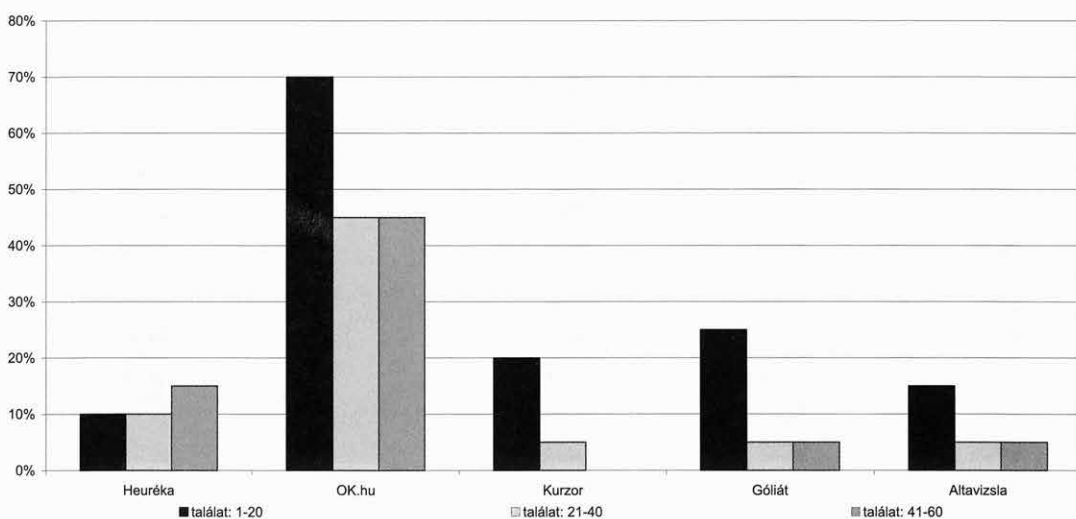
A „Rákóczi-szabadságharc” keresőkulcs találati pontossága

találatban; ezzel szemben az Origo-Vizslánál és a Góliátnál a nem monoton tendencia figyelhető meg. A Góliát és az Altavizsla között minimális különbség volt a releváns találatok számában, ami a második húsz találat esetében volt tapasztalható. Az Altavizsla azonos számú releváns találatot adott vissza a második és a harmadik húszas találatban. A Góliát nem szolgáltatott releváns találatot a második húsz találatban, azonban a harmadik húsz találatában megegyezett az Altavizsla ugyanezen a helyen lévő releváns találatainak a számával, de az Altavizsla monoton csökkenő tendenciát mutatott a releváns találatok lekérdezésében az első hatvan találatban.

2008 őszén végeztem egy kontroll-tesztet ugyanerre a keresőkulcsra vonatkozóan azzal a céllal, hogy tanulmányozzam a keresőknek milyen eltérő visszakeresési tendenciáik vannak a korábbi állapothoz képest. A tesztelés során kapott eredmények megtekinthetők a 4. diagramon. Megjegyzem, hogy a vizsgált keresők közül az Origo-Vizsla kereső 2006 decemberében megszűnt, ezért annak jelenlegi utódját, az OK.hu keresőt elemeztem, amely 2007 májusától áttért a Google keresőtechnológiára a korábban használt Yahoo helyett.

A 4-es diagramról elmondható, hogy a vizsgált keresők többsége az első húsz találatban keres-

te vissza a legtöbb releváns találatot, kivéve a Heuréka keresőt. Az OK.hu kereső releváns találatainak a száma kiemelkedően magas volt a többi vizsgált keresőéhez képest az első húsz találatban. A Góliát kereső második helyre került sokkal kevesebb releváns találatszámmal az első húsz találatban, majd pedig utána következtek a Kurzor, az Altavizsla és a Heuréka keresők egymástól kisebb releváns találat különbségekkel lemaradva. Az OK.hu, a Góliát és az Altavizsla keresők azonos monoton csökkenő tendenciával adták a releváns találatokat az első hatvan találatban. Mindhárom kereső a legtöbb releváns találatot hozta az első húsz találatban, továbbá ugyanannyi releváns találatot szolgáltatott a második és a harmadik húsz találatban. A Góliát és az Altavizsla keresők között csupán az első húsz találatra kapott releváns találatok számában volt eltérés. A Heuréka kereső nem monoton tendenciát nyújtott az első hatvan találatban, amely kisebb növekedést mutatott a harmadik húsz találatban. Ez a tendencia a Heuréka kereső nem kielégítő működését jelzi, mert alapesetben a keresőktől a releváns találatok csökkenését várjuk a találati listában előrehaladva. A Kurzor kereső szigorúan monoton csökkenő tendenciával szolgáltatva az első hatvanból kapott releváns találatokat. Annak tendenciájában nem jelentkezett



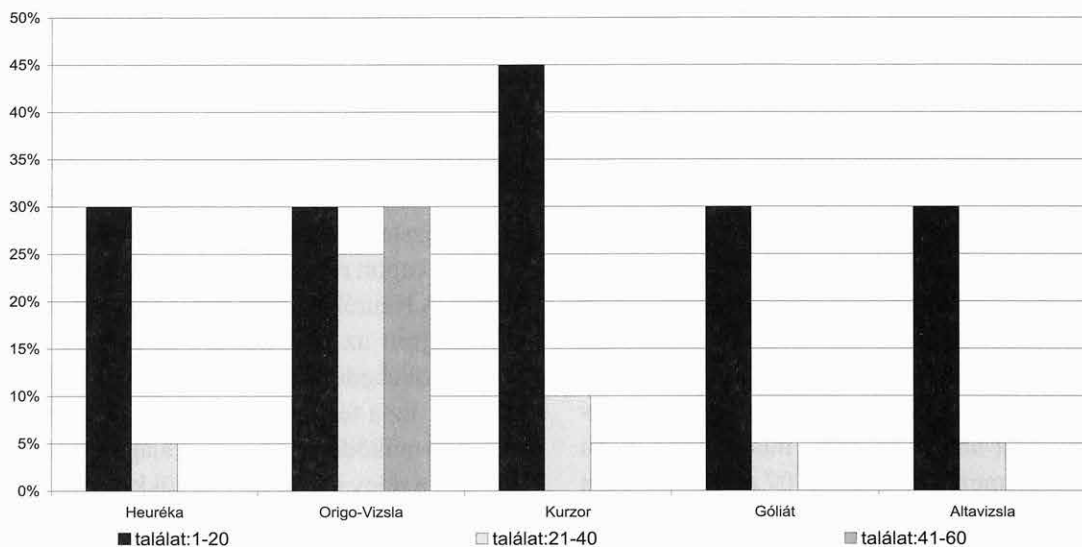
4. diagram

A „Rákóczi-szabadságharc” keresőkulcs – találati pontossága a kontroll-teszt során

változás a 2006-os állapothoz képest. Azonban az első húsz találatban hozott releváns találatoknak a száma erőteljesen visszaesett a korábbi állapothoz viszonyítva. Az Altavizsla kereső szintén ugyanazzal a monoton csökkenő tendenciával kereste vissza a releváns találatokat az első hatvan találatban, mint korábban. Továbbá megállapítható, hogy a többi vizsgált kereső tendenciája megváltozott az első hatvan találatban a korábbi állapothoz képest. Érdekes adat lehet

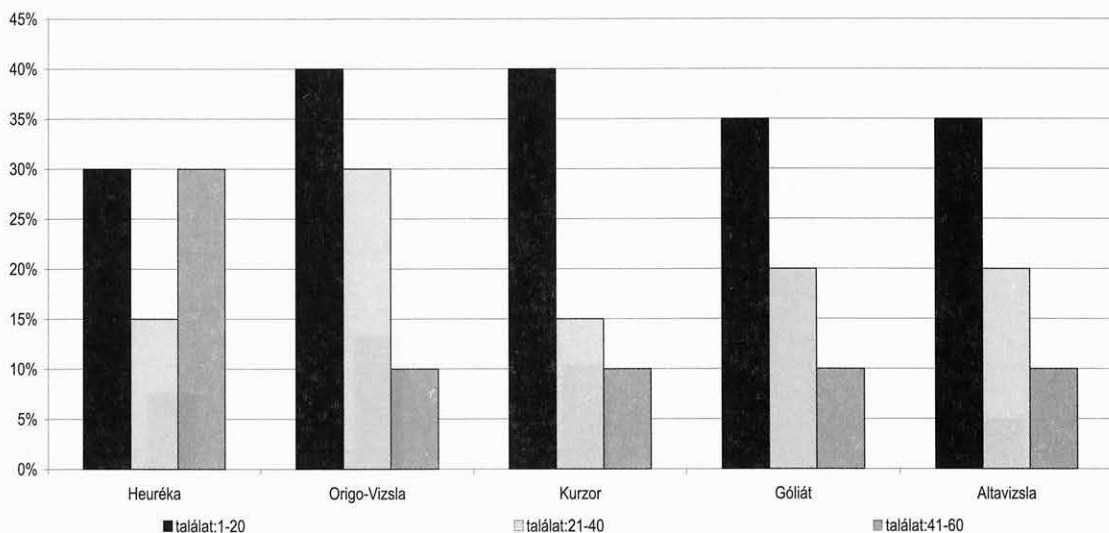
számunkra, hogy a tesztelés során összesen 25 új releváns dokumentumot kutattak fel a keresők a megadott keresőkérdésre, ami adatbázisaik tartalmának változását tükrözi.

Az 5. diagram alapján megállapítható, hogy az evészavar témájában a keresők első húsz találatában volt a legtöbb releváns találat. Ez alól az Origo-Vizsla volt kivétel, mert ugyanannyi releváns találatot adott az első húszban, mint a harmadikban. A Kurzor nyújtotta a legtöbb re-



5. diagram

Az „evészavar” keresőkulcs – találati pontossága



6. diagram

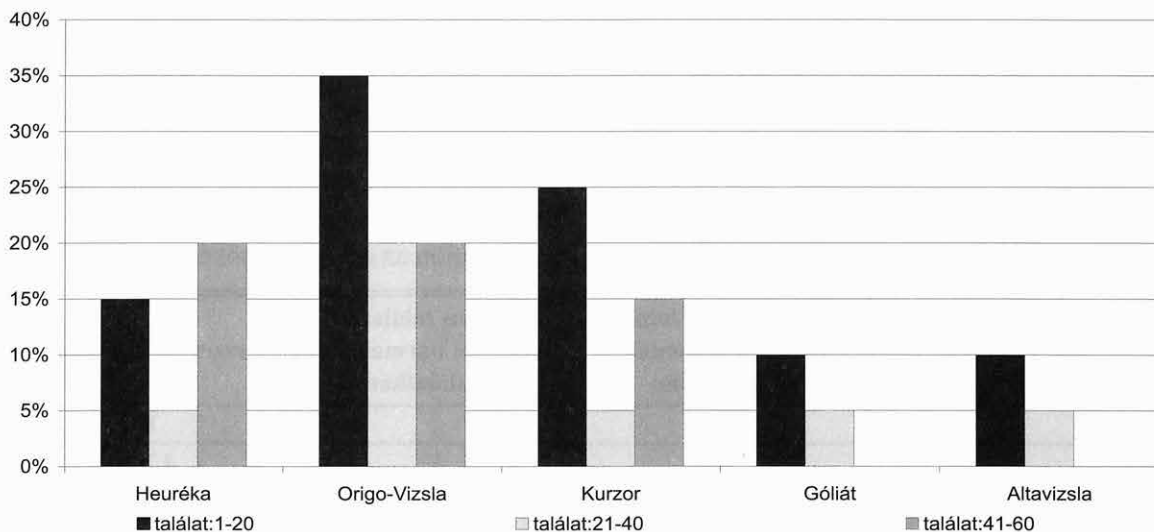
A „kegyhely” keresőkulcs találati pontossága

leváns találatot az első húsz találatban, a többi pedig azonos számú releváns oldalt hozott. Az Origo-Vizsla kivételével az összes többi keresőnél ugyanaz a szigorúan monoton csökkenő tendencia volt jellemző a releváns találatok lekérdezésében az első hatvan találatban. Ezen keresők közül egyik sem keresett vissza egyetlen releváns találatot sem a harmadik húsz találatban. Az Origo-Vizslánál viszont nem monoton tendencia volt megfigyelhető az első hatvanból származó releváns találatban.

A 6-os diagramról leolvasható, hogy az első húsz találatban erre a keresőkulcsra adták a keresők a legtöbb releváns találatot. A Heuréka eltérően viselkedett a többihez képest, mert ugyanannyi

első hatvan találat között. Azonban a Heuréka nem monoton tendenciát valósított meg az első hatvanból származó releváns találatban.

A 7-es diagramból látható, hogy erre a keresőkulcsra szinte mindegyik a legtöbb releváns találatot kereste vissza az első húsz találatban. A Heuréka kivétel, mert ott több releváns találatot szolgáltatott a harmadik húsz között, mint az első húsz találatban. Az Origo-Vizsla kutatta fel a legtöbb releváns találatot az első húsz találatban. A Kurzor kereső alacsonyabb releváns találatszámmal követte ugyanezen a helyen. A Heuréka került a harmadik helyre a releváns találatok száma alapján az első húsz találatban. A Góliát és az Altavizsla lettek az utolsók és ezek



7. diagram

A „szökőár” keresőkulcs találati pontossága

releváns találatot adott az első húsz találatban, mint a harmadik húszban. Az Origo-Vizsla és a Kurzor esetében magas és azonos számú a releváns találatok száma az első húsz találatban. A Góliát és az Altavizsla azonos számú releváns találatot szolgáltatott az első húsz találatban, ezért együtt kerültek a rangsor második helyére. A Heuréka kereső kevéssel ugyan, de lemaradt az előző két kereső releváns találataitól az első húsz találatban. A Heurékát leszámítva a többinél megfigyelhető egy szigorúan monoton csökkenő tendencia a releváns találatok lekérdezésében az

adtak le azonos számú releváns találatot is az első húszban. A Góliát és az Altavizsla keresőknél szigorúan monoton csökkenő tendencia jelentkezett a releváns találatok lekérdezésében az első hatvan találatban, ráadásul azonos számú releváns találatot nyújtottak az első hatvan találatban erre a keresőkulcsra. Az Origo-Vizsla monoton csökkenő tendenciát mutatott a releváns találatok lekérdezésében az első hatvan találatban, míg a Heuréka és a Kurzor keresők nem monoton tendenciával jellemezhetők.

Mi következik az adatokból?

A diagramok azt mutatják, hogy a vizsgált magyar nyelvű keresők többsége arra törekedett, hogy a legtöbb releváns találat az első húsz találatban jelenjen meg mind a négy keresőkulcs esetében. A Kurzor az első két keresőkulcs értékre a legtöbb releváns dokumentumot adta az első húsz találatban, míg a harmadik (<kegyhely>) keresőkulcsra kapott találatsszámmal holtversenybe került az Origo-Vizslával, mivel mindketten azonos számú releváns találatot szolgáltatott ugyanazon a vizsgált helyen. A negyedik (<szökőár>) keresőkulcs esetében azonban a Kurzor már elveszítette az első helyét, mert megelőzte az Origo-Vizsla a visszakeresett releváns dokumentumok száma terén az első húsz találatban. Amennyiben ezeket az eredményeket a felhasználó szempontjából közelítjük meg, akkor megállapíthatjuk, hogy a Kurzor és az Origo-Vizsla bizonyult a legjobbnak a használat során, hiszen ezeknél volt a legtöbb releváns oldal az első húsz találatban. (Megjegyzem, hogy a Góliát és az Altavizsla keresők minimális el-

téréssel ugyan, de közel azonos releváns találatmennyiséget kerestek vissza az első hatvan találatban mind a négy keresőkulcsra, továbbá azok többnyire – egy keresőkérdés kivételével – megegyező tendenciával hozták a releváns dokumentumokat az első hatvan találatban mind a négy keresőkulcs értékre. E két keresőnek a keresőkulcsokra visszakeresett összes találatainak száma viszont különböző volt, ezért komolyabb következtetés nem vonható le az adatbázisaik egyezésére vonatkozóan.)

A találatok pontossága

A következő lépésben az első, a második, valamint a harmadik húsz találatban kapott releváns találatokat hasonlítottam össze a négy kérdést tekintve. Figyelembe vettem, hogy a keresők külön-külön összesen hány találatot keresnek vissza az egyes keresőkulcsokra. A 2. táblázatban összefoglalom az első, a második, és a harmadik húsz találatban lévő releváns találatok számát, valamint az összes találat számát, amit a keresők külön-külön az egyes keresőkérdésekre adtak.

Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
5	3	2	836
7	3	5	4 240
11	3	2	157
6	0	1	16 641
6	1	1	60 894
6	1	0	125
6	5	6	1 070
9	2	0	1 422
6	1	0	5 296
6	1	0	5 296
3	1	4	446
7	4	4	6 560
5	1	3	12 766
2	1	0	247
2	1	0	247
6	3	6	255

Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
8	6	2	2 790
8	3	2	3 469
7	4	2	5 286
7	4	2	5 286
<i>Összesen:</i> 123	48	42	133 329

2. táblázat

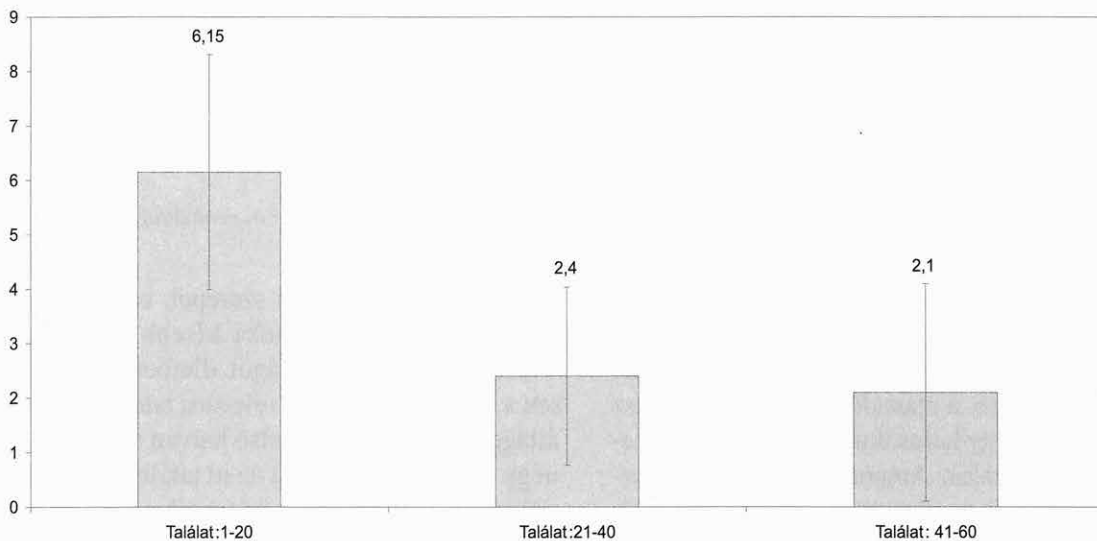
A keresőkulcsokra lekérdezett összes találatszámok és a releváns találatszámok

Az adatok alapján nincs konkrét összefüggés a releváns találatok száma és az összes találat száma között. Ez többek között abban is nyilvánul, hogyha például közel azonos nagyságrendű összes találatszámokat veszek alapul, akkor nem kapok egymáshoz hasonló nagyságrendű releváns találatszámokat, csupán a releváns találatszámoknak különféle variációit nyerem. A táblázatrol leolvasható, hogy a keresők összesen hat esetben egyetlen releváns találatot sem adtak a harmadik húsz találatban a megadott keresőkulcsokra. Ezért ezekben az esetekben nem határozható meg meggyőző következtetés a releváns találatok számára vonatkozóan a találathalmaz további részére.

A szórás és a várható érték

A szórást és a várható értéket különböző megközelítésekből mutatom be, és diagramokkal illusztrálom. Az oszlopdiagramok felett közlöm a várható értéket, függőleges vonallal jelölve a tapasztalati szórást és a mintaátlag szórását a pozitív és negatív értéktartományban. A diagramok elemzésével kapcsolatos megjegyzések a minta nagysága miatt természetesen nem tekinthetők statisztikai következtetéseknek, azonban határozottan kijelölhetik egy jövőbeli statisztikai elemzés célkitűzéseit.

A 8. diagram adatai szerint a vizsgált keresők releváns találatainak átlagértéke mind a négy



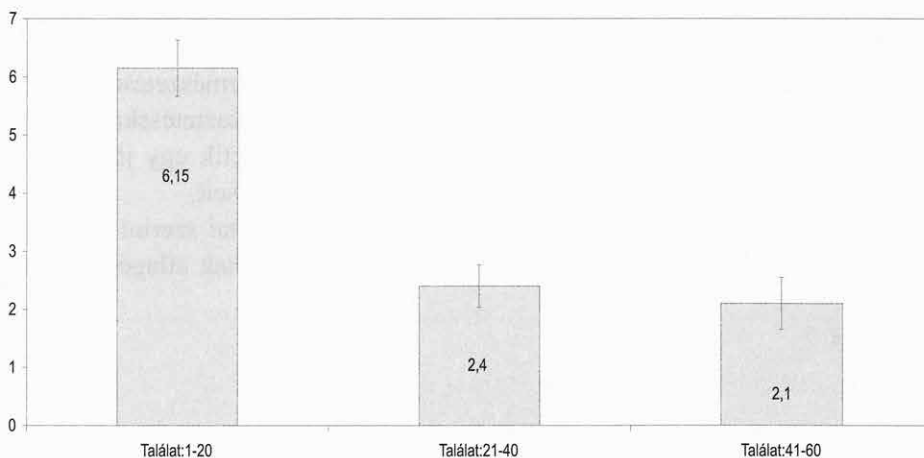
8. diagram

Az első, a második és a harmadik húsz találatban lévő releváns találatok átlaga és tapasztalati szórása a keresőkulcsokra és a keresőkre összesítve

keresőkulcsra nézve szigorúan monoton csökkenő. A releváns találatok tapasztalati szórása az első húsz találatban azt jelzi, hogy a valószínűségi változó értéke nagyobb valószínűséggel tér el az elméleti várható értéktől, mint a második és a harmadik húsz találatban jelentkező tapasztalati szórás esetében. A vizsgált keresők hatékonysága az első húsz találatban mind a négy keresőkulcsra a legnagyobbak tekinthető, mivel releváns találataik átlagértéke itt bizonyult a legmagasabbnak. Alacsony a keresők hatékonysága a második és a harmadik húsz találatban mind a négy vizsgált keresőkérdésre, amit a releváns találatok alacsony átlagértékei tükröznek. Minimális különbség figyelhető meg a keresők hatékonyságában a második és a harmadik húsz találatban is, mert ezeken a helyeken csekély eltérés mutatkozott a releváns találatok átlagértékei között.

mintaátlag hibáját mutatja számunkra.

A 10-es diagram alapján elmondható az, hogy a Kurzor esetében a legnagyobb a valószínűsége, hogy a valószínűségi változó értéke eltér az elméleti várható értéktől az első hatvan találatban mind a négy keresőkulcsra, de a valószínűség nagy mértékben csökken a Góliát és az Altavizsla keresőnél az első hatvan találatban a négy keresőkérdésre, az pedig kisebb mértékben ugyan, de tovább csökken a Heuréka keresőnél az első hatvan találatban mind a négy keresőkulcsra. Az Origo-Vizslánál a legkisebb a valószínűsége annak, hogy a valószínűségi változó értéke nagy mértékben eltérjen az elméleti várható értéktől. Ez a kereső tűnik a leghatékonyabbnak az első hatvan találatban mind a négy keresőkulcsra, hiszen releváns találatainak átlagértéke a legmagasabb a többi keresőéhez viszonyítva. Ugyanebből a szempontból a He-

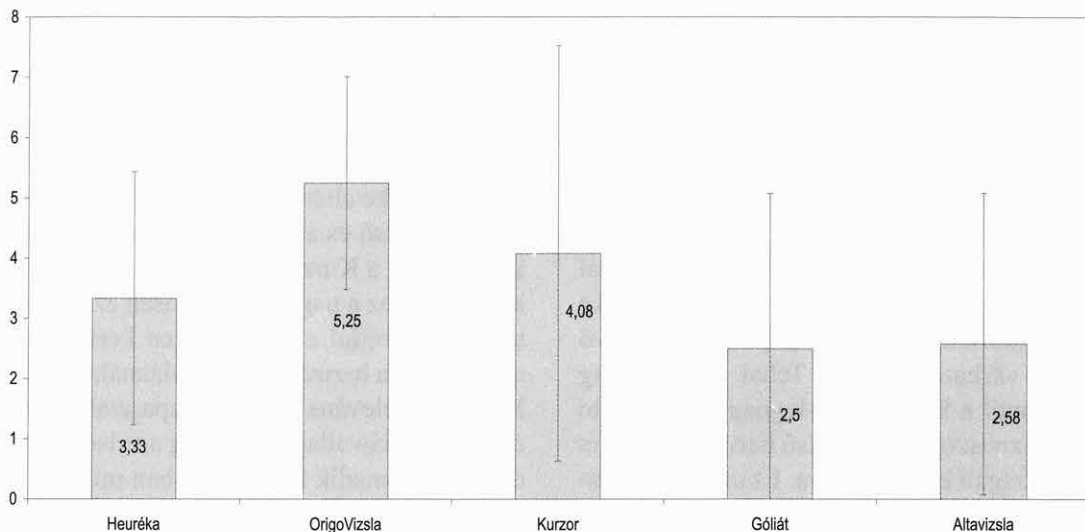


9. diagram

Az első, a második és a harmadik húsz találatban lévő releváns találatok átlaga és mintaátlag szórása a keresőkulcsokra és a keresőkre összesítve

A 9-es diagramról leolvasható, hogy egy véletlenszerűen választott minta átlagértéke csekély mértékben tér el a populáció elméleti várható értékétől az első, a második és a harmadik húsz találatban a négy keresőkulcsra és az összes keresőre vonatkozóan. Amennyiben a mintaválasztást, a mérést és az átlagszámítást többször is elvégezzük, akkor a kiszámított mintaátlagok csak minimális mértékben fognak eltérni a populáció elméleti várható értékétől, tehát ez a diagram a

uréka a második helyen szerepel, bár releváns találatainak az átlagértéke kisebb az Origo-Vizslánál. A hatékonyságot illetően a Kurzor lett a harmadik, hiszen a releváns találatainak az átlagértéke alacsony az első hatvan találatban a négy keresőkulcsra, amit az itt található rendkívül magas tapasztalati szórás érték eredményez. A keresés hatékonysága szempontjából az utolsó helyre a Góliát és az Altavizsla került. (Meg kell jegyezni, hogy nincs eltérés e két kereső haté-



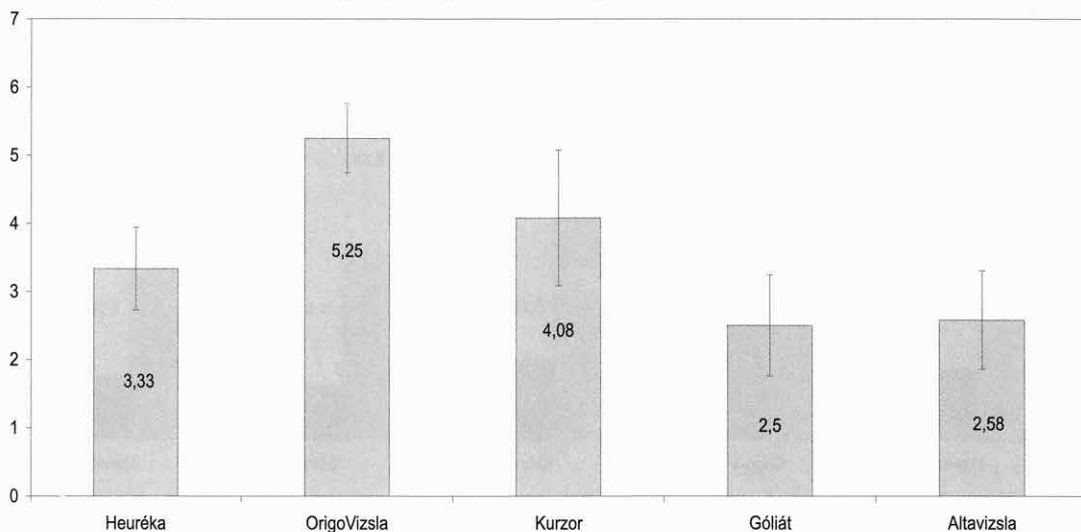
10. diagram

Az első hatvan találatban mért releváns találatok átlaga és tapasztalati szórása a négy keresőkulcsra összesítve keresőnként

konszága között, mivel azok tapasztalati szórása közel azonos nagyságrendű, valamint minimális különbség jelentkezett releváns találatok átlagértékei között az első hatvan találatban mind a négy keresőkulcsra.)

Feltételezésem szerint a diagramon látható számottevő szórás értékek abból adódhatnak, hogy a keresőknél az új weboldalak indexelésének az átfutási ideje meglehetősen lassú, ez pedig kihat

a keresők hatékonyságára. Tudomásom szerint a Heuréka és az Origo-Vizsla keresőknél a keresőrobot átlagosan 2–3 hetente gyűjti be az új oldalakat, míg a Kurzornál néhány hét után frissül az adatbázis (Zádori, 2007). A Heuréka és a Kurzor a híryanagokat napi rendszerességgel követik nyomon. Hiteles információk (Hódi, 2007) szerint a jelenleg publikusan elérhető Góliát egyáltalán nem frissíti adatbázisát. Az Altavizsla



11. diagram

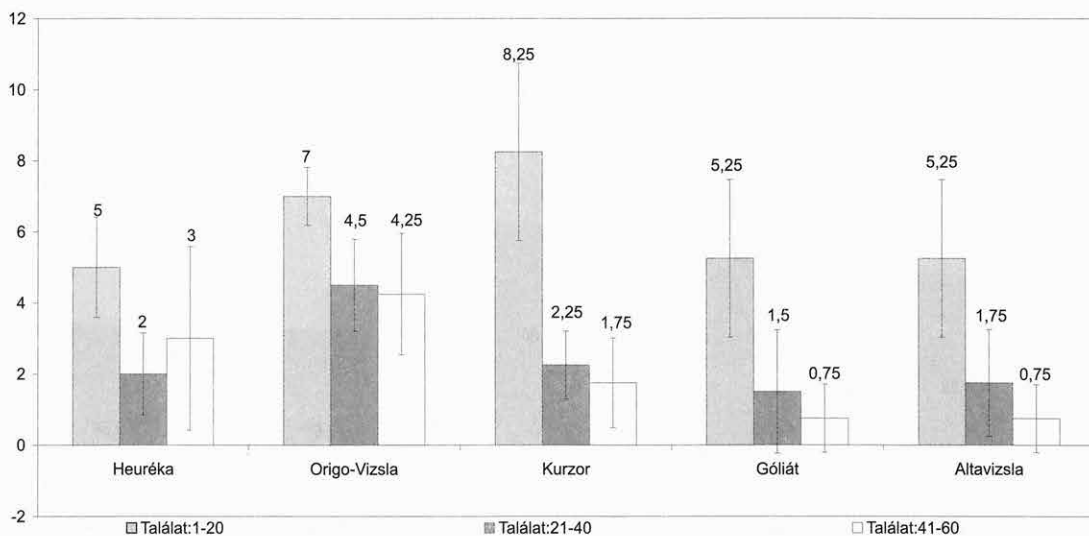
Az első hatvan találatban mért releváns találatok átlaga és mintaátlag szórása a négy keresőkulcsra összesítve keresőnként

(Bocsán 2007) a Góliát keresőt használja a háttérben, saját indexet külön nem hoz létre. Ez utóbbi értesülést támasztják alá a vizsgálat eredményei is, mert a Góliát és az Altavizsla kereső egyetlen weboldal kivételével ugyanazokat a releváns dokumentumokat keresték vissza mind az öt témakörre vonatkozóan.

A 11-es diagram azt tükrözi, hogy a Kurzornál egy tetszőlegesen választott minta átlagértéke jelentős mértékben eltér a populációban lévő elméleti várható értéktől. Tehát a mintaátlag hibája ennél a keresőnél a legnagyobb a többi keresőhöz viszonyítva az első hatvan találatban a négy vizsgált keresőkulcsra. Ez után következnek a Góliát és az Altavizsla kicsit kisebb, de közel azonos nagyságrendű mintaátlag hibával az első hatvan találatban a négy keresőkérdésre. A Heuréka és az Origo-Vizsla keresőknél a mintaátlag-hibája tovább csökken ugyanezen a vizsgált helyen a négy keresőkulcsra.

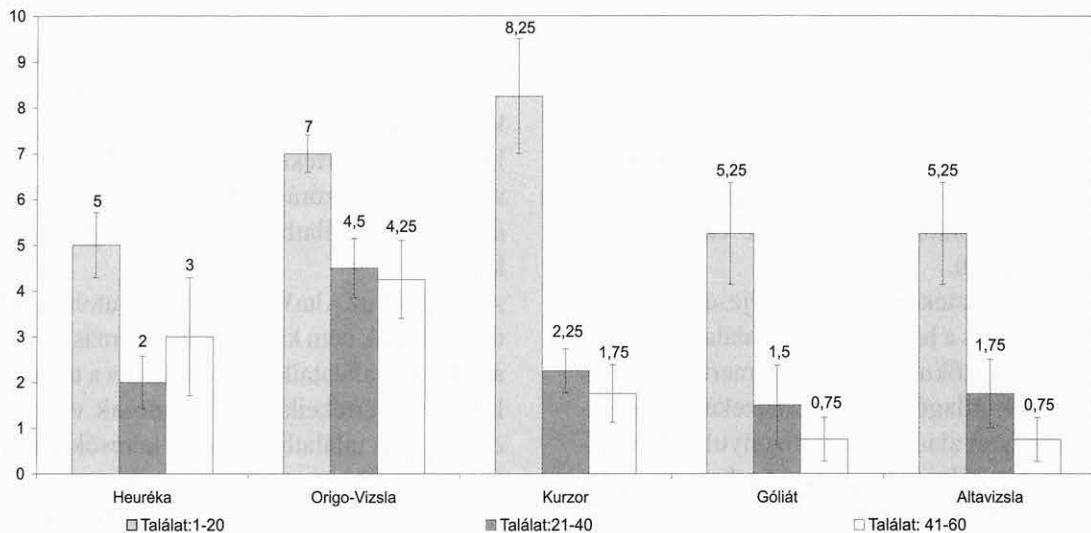
A 12. diagramról elmondható, hogy a Heuréka kereső kivételével az összes többi keresőnél a releváns találatok átlagértékei szigorúan monoton csökkenő tendenciát követnek az első, a második és a harmadik húsz találatban a négy keresőkulcsra. A Heuréka keresőnél a releváns találatok átlagértékei nem monoton tendenci-

át mutatnak az első, a második és a harmadik húsz találatban a négy vizsgált keresőkérdésre, továbbá a Heuréka és az Origo-Vizsla keresőknél a harmadik húsz találatban a legnagyobb a valószínűsége annak, hogy a valószínűségi változó értéke eltér az elméleti várható értéktől e keresők első és a második húszas találatához képest. Míg a Kurzor, a Góliát és az Altavizsla keresőknél ez a nagy valószínűség az első húsz találatban fordul elő ugyanezen keresőknek a második és a harmadik húsz találatához képest. Mindez a releváns találatok tapasztalati szórás értékei alapján állapítható meg az első, a második és a harmadik húsz találatban mind a négy keresőkulcsra. Jól látható a diagramon az is, hogy a keresők az első húsz találatban arra törekcszenek, hogy a leghatékonyabbak legyenek a négy keresőkulcsra. Ez azzal magyarázható, hogy mindegyikőjük releváns találatainak az átlagértéke sokkal magasabb az első húsz találatban, mint a második és a harmadik húsz találatban. A Góliát és az Altavizsla között nincs eltérés a hatékonyság tekintetében, hiszen a releváns találataik közel azonos átlagértéket és tapasztalati szórás értéket tükröznek az első, a második és a harmadik húsz találatban mind a négy keresőkérdésre.



12. diagram

A keresők első, második és harmadik húsz találatában lévő releváns találatok átlaga és tapasztalati szórása a négy vizsgált keresőkulcsra



13. diagram

A keresők első, második és harmadik húsz találatában lévő releváns találatok átlaga és mintaátlag szórása a négy keresőkulcsra

A diagram azt mutatja, hogy a Heuréka és az Origo-Vizsla keresőknél a harmadik húsz találatban egy tetszőlegesen választott minta átlagértéke nagyobb mértékben tér el a populáció elméleti várható értékétől ugyanezen keresőknek az első és a második húsz találatához képest. Ez a jelentősebb mértékű eltérés figyelhető meg a Kurzor, a Góliát és az Altavizsla keresők esetében az első húsz találatban ugyanezen keresőknek a második és a harmadik húsz találatához képest. Ezenkívül a Góliát és az Altavizsla keresőknél nagyon hasonló mértékű eltérésről beszélhetünk egy tetszőlegesen választott minta átlagértéke és a populáció elméleti várható értéke között az első, a második és a harmadik húsz találatban, tehát a mintaátlag hibája közel azonos nagyságrendűnek tekinthető ennél a két keresőnél.

Összefoglalás

Az adatokból megállapítható, hogy a pontosság megváltozik valamelyest az egyes vizsgált keresők esetében, tehát levonható valamilyen következtetés a pontosság alakulására vonatkozóan, ami érvényes a találatalmaz egészére is. A pontosság csak minimális mértékben változott

a Góliát és az Altavizsla keresőknél, mert azok közel azonos számú releváns találatot kerestek vissza az első, a második és a harmadik húsz találatban mind a négy keresőkérdésre. (Ld. 3., 4., 5., 6., 7. diagramokat.)

A vizsgálat során nyert adatokból kiderült, hogy nincs evidencia szintű összefüggés az első hatvan találatban lévő releváns találatok száma és a keresőkulcsokra kapott összes találat száma között. Ebből adódóan nem határozható meg meggyőző következtetés a releváns találatok számára vonatkozóan a találatalmaz hátralévő részében. Tehát ezt a hipotézist el kellett vetnünk. (Ld. a 2. táblázatot.)

– A találati lista elején lévő találatok pontosak, ezután fokozatos csökkenő értéket mutatnak.

A vizsgált keresők többsége arra törekedett, hogy a legnagyobb mennyiségű releváns találatot az első húsz találatban a négy keresőkulcsra szolgáltassa.

Az első hatvan találatban háromféle tendenciában keresték vissza a releváns találatokat, amelyek a következők voltak: szigorúan monoton csökkenő, monoton csökkenő és nem monoton. Ebben az esetben a keresőket a használat szempontjából vizsgáltam, ahol kizárólag az első, a második és a harmadik húsz

találatban visszakeresett releváns találatok mennyiségét vettem figyelembe. (Ld. a 3., 4., 5., 6., 7. diagramokat.)

A hatékonyság szempontjából mind a négy keresőkérdés esetében az első húsz találatban bizonyultak a keresők a legjobbnak, mivel releváns találataik átlagértéke ekkor volt a legmagasabb.

A keresők hatékonysága erőteljesen csökkent a második és a harmadik húsz találatban (mind a négy keresőkulcs esetében), mert a releváns találataik átlagértéke ugyanezekben a helyeken nagyon alacsonynak bizonyult. Minimális eltérés volt megfigyelhető a keresők hatékonyságában a második és a harmadik húsz találatban (mind a négy keresőkulcsra), mert csekély különbség volt a releváns találataik átlagértékei között ezeken a vizsgált helyeken. (Ld. a 8. diagramot.)

– *A vizsgálat eredményeként kiderül, melyek azok a magyar nyelvű keresők, amelyek a legpontosabb találatokat szolgáltatják.*

Használat szempontjából a Kurzor és az Origo-Vizsla minősült a legjobbnak, mert mindkét kereső arra törekedett, hogy a lefuttatott keresésekre a lehető legtöbb releváns találatot kutassa fel az első húsz találatban (ld. a 3., 5., 6., 7. diagramot.). Amennyiben a keresők találatsszolgáltatásának az egyenletes minőségére helyezük a hangsúlyt az első hatvan találatban (és nem csupán a visszakeresett releváns találataik mennyiségére fókuszálva), akkor ez utóbbtól eltérő eredményeket kapunk. (Megjegyzem, hogy egy kereső találatsszolgáltatásának az egyenletes minőségét a hatékonyság aspektusa fejezi ki számunkra, amit a visszakeresett releváns találatok átlagértékei és tapasztalati szórás értékei figyelembevételével kapunk.)

Az Origo-Vizsla kereső bizonyult a leghatékonyabbnak az első hatvan találatban mind a négy keresőkulcsra, mert a többi vizsgált keresőhöz képest a releváns találatainak az átlagértéke a legnagyobb volt. A hatékonyságot tekintve a Heuréka került a második helyre, bár releváns találatainak az átlagértéke sokkal

kisebb volt az Origo-Vizslához képest az első hatvan találatban a vizsgált négy keresőkérdésre. Ugyanebből a szempontból a Kurzor került a harmadik helyre, mert releváns találatainak átlagértéke alacsonynak bizonyult – a tapasztalati szórás figyelembevételével – az első hatvan találatban a vizsgált négy keresőkulcsra.

A Góliát és az AltaVizsla lettek az utolsók. Hatékonyságuk nem különbözött egymástól, mert a releváns találataik átlagértékei és a tapasztalati szórás értékeik közel azonosak voltak az első hatvan találatban a négy keresőkérdésre. (Ld. a 10. diagramot.)

– *A vizsgálat eredményeként értékelhető, hogy a magyar nyelvű keresők milyen mértékben képesek megtalálni a releváns találatok halmazát, azaz találathalmazaik milyen arányban tartalmazzák azokat.*

A relatív teljességre kapott százaléértékek alapján a következő rangsort állítottam fel: az Origo-Vizsla találta meg a vizsgált releváns találathalmaz legtöbb dokumentumát, mert azok 42,85%-át fedezte fel a weben. A második helyre a Kurzor, a Góliát és az Altavizsla kerültek 23,80%-kal, hiszen azonos mennyiségű releváns dokumentumot kerestek vissza. A Heuréka kissé lemaradt az előző három kereső visszakeresési teljesítményétől, mert az a találathalmazban lévő releváns dokumentumoknak csak 19,04%-át találta meg a weben. (Ld. a 2. diagramot.) Ehhez kapcsolódóan még egy fontos mutatószámra hívnám fel a figyelmet, mely szerint a magyar nyelvű keresők átlagosan 5,6 releváns dokumentumot találtak meg a weben <csiperkegomba> témakörrel kapcsolatban, ami a vizsgált halmazban lévő huszonegy találat 26,66%-ának felelt meg. Tehát a magyar nyelvű keresők a releváns találatok halmazának alig több mint egynegyedét keresték vissza átlagosan a <csiperkegomba> témakörben, ami meglehetősen szerény teljesítménynek számít.

Megállapítható, hogy további vizsgálatoknak kell alátámasztaniuk ezt a kapott eredményt annak érdekében, hogy megalapozottabb

következtetésre jussunk a magyar nyelvű keresők hatékonyságával kapcsolatban. Azt jegyzem csak meg, hogy demonstráláson kívül a szándékom az volt, hogy ezek a kérdések elemezhetőek. Vizsgálatomban elsősorban egy tesztkulcssor összeállítására törekedtem, amely óvatosan bővíthető egy részletesebb statisztikai vizsgálat számára. További kutatási feladat lehet számunkra a hatékonyságkutatási esetek gazdag és komplex gyűjteményének létrehozása az általunk kialakított modellbe illeszkedve, amely alapját képezhetné egy mélyebb szintű statisztikai vizsgálatnak.

<http://infolab.stanford.edu/pub/papers/google.pdf>
(2007. márc. 2.)

A Heuréka tájékoztató segédlete
<http://www.heureka.hu/heureka/tippek.html> (2007. márc. 9.)

HÓDI Imre [Személyes közlés] (2007. aug. 13.)

HORVÁTH T.–SÜTHEŐ P.: A tartalmi feltárás. In: Könyvtárosok kézikönyve. 2. köt. Feltárás és visszakeresés. Szerk. Horváth T.–Papp I. Bp.: Osiris, 2001. 35–186. p.

OPPENHEIM, C.– MORRIS, A.– MCKNIGHT, C.– LOWLEY, S.: The evaluation of WWW search engines. = Journal of Documentation, 56. vol. 2000. 2. no. 190–211. p.

Az Origo-Vizsla tájékoztató segédlete
http://www.vizsla24.hu/pages/sugo_tartalomjegyzek.html
(2005. aug. 11.)

UNGVÁRY R.: Az információkeresés értékelése. In.: Osztályozás és információkeresés: kommentált szöveggyűjtemény. 2. köt. Az információkeresés és elmélete. Szerk. Ungváry R., Orbán É. Bp.: OSZK, 2001. 535 p.
<http://mek.oszk.hu/01600/01683/pdf/01683-2.pdf> (2007.11.17.)

UNGVÁRY R. – VAJDA E.: Könyvtári információkeresés. Bp. Typotex, 2002. 169 p.

Taxaurusz: Növény- és állatrendszertani fogalmak makrotezaurusza: 3.0 változat. [összeáll.] Ungváry R.; [közr. az] Országos Széchényi Könyvtár. Bp.: OSZK, 2000. XVII, 4, 16, 98 p.
<http://mek.oszk.hu/00000/00077> (2007. ápr. 30.)

ZÁDORI Janka [Személyes közlés] (2007. aug. 10.)

Irodalom

ALBERT L.–LOCSMÁNDI Cs. – VASAS G.: Ismerjük fel a gombákat! Gabó, 1995. 191 p.

BERNERS-LEE, T. – FIELDING, R. – FRYSTYK, H.: Hypertext Transfer Protocol–HTTP/1.0. May 1996.
<http://www.rfc-editor.org> (2007.08.10.)

BOCSÁN Róbert [Személyes közlés] (2007.szept. 4.)

BRIN, S.– PAGE, L.: The anatomy of a large-scale hypertextual web search engine. = Computer Networks and ISDN Systems, 30. vol. 1998. 1–7. no. 107–117. p.

1. függelék

Csiperkegomba – releváns találatok

1. Mellékletek I.: melléklet. FFV-24 számú ENSZ/EGB szabvány. (UN/ECE STANDARD FFV-24) Termesztett csiperkegomba. (Ajánlott nemzetközi kereskedelmi szabvány) 1. ENSZ/EGB szabványok friss gyümölcsre és zöldségre ... Ez a szabvány a friss állapotban forgalomba kerülő fogyasztásra szánt *Agaricus bisporus* termesztett fajtáinak termőtestére vonatkozik, és azokat a követelményeket határozza meg, amelyek az értékesítésre előkészített és csomagolt csiperkegombára vonatkoznak. Nem tárgyala

szabványnak az ipari felhasználásra szánt csiperkegomba. ...

<http://www.omgk.hu/MGUT6/mel1.html> (Origo-Vizsla, Kurzor, Góliát, Altavizsla)

2. Dr. Rác László: Talaj és termesztett csiperkegomba fő- és nyomelemmigrációjának vizsgálata. Doktori értekezés összefoglalója. Debreceni Egyetem Kémiai Tudományok Doktori Iskolája, Debrecen 1998.

... Vizsgáltam a zárt térben termesztett csiperkegomba fejlődését, illetve a termésideszak alatt annak nyomelemtartalom-változását, amikor a

- termőtalajba, illetve az azt fedő felső rétegbe (ún. takarótalajba) ismert és változó mennyiségben juttattuk be a gomba fő elemeit (K, Na, Mg, Ca), az esszenciális nyomelemeket (Mn, Zn, Cu, Co, Ni), valamint a toxikus nyomelemeket (Cd, Cr, Hg, Pb). Az analitikai vizsgálatokat elsősorban azért tartjuk fontosnak, mert a környezetből (levegőből, vízből és talajból) egyre gyakrabban juthatnak mérgező nyomelemek ezen élelmiszerekkel szervezetünkbe ...
<http://www.doktorijegyzek.hu/jeg.php?subid=1&pid=228&back=3> (Heuréka, Góliát, Altavizsla)
3. Dr. Rác László: Makro- és mikroelemek felszívódása a talajból és azok hatása néhány termesztett gombafajtára. Doktori (Ph.D.) értekezés tézisei.
 ... A kontroll- és fémsóoldatokkal kezelt táptalajon termelt csiperkegomba mintavétele nagyrészt a négy terméshullásban történt. Általában 3-3, illetve 5-5, egyes esetekben 6-6 párhuzamos termesztőzsákból megfelelő számú 2, 3, illetve 5 gombamintát szedtünk. Műanyag késsel szűrőpapíron negyedelve, esetleg nyolcadolva átlagmintavétel alapján 20 g körüli mennyiséget mértünk be analitikai mérlegen ...
<http://kemia.ektf.hu/racz.htm> (Heuréka, Kurzor)
4. Csiperkegomba: környezeti igénye, a termesztés feltételei, a termesztés módszerei. In: Zöldségtermesztők kézikönyve/ szerk. Balázs Sándor. Bp.: Mezőgazda Kiadó, 2004.
 ... Hő. A csiperkegomba hőigénye az egyes termesztési ciklusokban eltérő: más az átszövetési (1) időszakban, más a lappangás (2) idején és ismét más a termesztési (3) időszakban. Az átszövetési időszakában 22–25°C-ot igényel optimális fejlődéséhez. ...
<http://www.hik.hu/index.asp?a=r&r=90%2C261&oid=57> (Kempelen Farkas Hallgatói Információs Központ, Kempelen Farkas Digitális Tankönyvtár honlapja) (Origo-Vizsla)
5. Kétspórás csiperke – *Agaricus bisporus*
 ... Kalap: 5-10 cm átmérőjű; fiatalon félgömb alakú, domború, majd ellaposodik, végül benyomottá válhat; fehéres, piszkos-, sötétbarna, széle világosabb; felületét barnás, lenyomott, sugarasan szálás pikkelyek díszítik; széle sokáig begöngyölt, gyapjas szálak, pelyhek lóghatnak le róla. Lemezek: sűrűn és szabadon állók, először rózsaszínűek, majd feketésbarnára sötétednek; éle fehéresen csipkézett. ...
<http://www.gombanet.hu/gomba.php?gid=186> (GombaNET.hu – a Magyar Mikológiai Társaság honlapja) (Heuréka, Origo-Vizsla)
6. Szarvas József: A bazídiumos gombák szaporodása, élelciklusa II.: a kétspórás csiperke (*Agaricus bisporus*) = Magyar Gombahíradó, (2003)
 ... A kétspórás csiperke fejlődésmenetének ismerete nélkülözhetetlen a nemesítői munkához. A fejlődési ciklusa a legtöbb "bazídiumos nagygombához" képest eltérő. Szaporodásukat korábban heterotallikusnak vélték, azonban ma tudjuk, hogy másodlagosan homotallikus fajról van szó (Fritsche, 1991). A sejtei sokmagvúak, nincs egyértelmű különbség a monospórás és a multispórás eredetű micélium között és mind a monospórás mind a multispór micélium képes termőtestképzésre anélkül, hogy más hifával anasztomizálnának. ...
<http://free.x3.hu/korona/other/bazid02.pdf> (Heuréka, Origo-Vizsla)
7. Kocsner Nóra: A gombaszúnyogok elleni védekezés entomopatogén nematodákkal. TDK dolgozat összefoglalója. Veszprémi Egyetem, Georgicon Mezőgazdaságtudományi Kar Növényvédelmi Állattani Tanszék, Keszthely 2004.
 ... A termesztett csiperke védelme nehézkes, mivel termesztéséhez sok szerves anyagot tartalmazó és megfelelően érlelt komposzt szükséges, amely nemcsak a csiperkegombának, hanem más patogén mikroszervezetnek is táptalaja, illetve élelciklusát tekintve gyors lefutású szervezet, emiatt a vegyszeres védekezés a termesztés idején korlátozott. ...
<http://www.vein.hu/tdk/ITDK2004/AT/AT-KocsnerN.doc> (Origo-Vizsla, Kurzor)
8. Rovarpatogén fonálféreg fermentatív termelése, alkalmazása. Szerződésszám: OMFB-01766/00 (BIO-012/00).
 ... Elsősorban a csiperkegomba (*Agaricus bisporus*) termesztésében a Diptera, Sciaridae taxonokba tartozó *Lycoriella solani*, *L. melii* fajok lárvái a termőtestek megrágásával, elpusztításával főként a tavaszi, nyári és őszi időszakokban nagy termés kiesést okoznak. Magyarországon a

- gombalegyek ellen leggyakrabban használt szer a Dimilin 25WP (Diflubenzuron). Korlátozott azoknak a vegyszereknek a száma, melyeket a Sciarid-lárva ellen használhatunk, mivel a legyek egy része rezisztenssé válik az aktív hatóanyaggal szemben ...
<http://www.nkth.gov.hu/letolt/k+f/Biotech2000.doc> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
9. Rózsai Gábor: A blokkos csiperkegomba termesztéstechnológiája. Kecskeméti Főiskola Kertészeti Kar. Oktatási segédanyag (MSPPT prezentáció)
 ... Csiperkegomba termesztés története Magyarországon: az 1800-as évek második felétől már nagyobb mennyiségben termesztünk. 1926-ban dr. Makó László: első magyar szinttényeszeti gombacsíra. 1930-tól mind a csíragyártás, mind a gombatermesztés gyorsan fejlődik. 1930-ban 200.000 m²-en évi 600 tonna terem. A világon a harmadik helyen, Európában a második helyen áll Magyarország. ...
<http://iatt.ttk.pte.hu/> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
10. Kétspórás csiperke. Familia: agaricaceae. Család: csiperkefélék
 ... Kalap: 4-10-(12) cm, gömb vagy félgömb alakból domború, ritkán kiterülő; széle sokáig begöngyölt, túlnő a lemezeken; felszíne változatos: fiatalon sima, később vagy sima, vagy kisebb-nagyobb koncentrikus pikkelyekkel; színe változatos, a fehértől a világosbarnáig vagy kakaóbarnáig. Lemezek: keskenyek, sűrűk, rózsaszínek, öregen bíborbarnák, szabadok. ...
http://www.terra.hu/gomba/html/agaricus_bi_sporus.html (Origo-Vizsla, Kurzor, Góliát, Altavizsla)
11. Mészáros Gábor Sándor: Csiperkegomba záskos termesztés-technológiájának értékelése, különös tekintettel a minőségbiztosításra. Szakdolgozat. Veszprémi Egyetem, Georgicon Mezőgazdaságtudományi Kar Növénytan és Növényélettani Tanszék, Keszthely 2001.
 ... A kétspórás csiperke a természetben leginkább trágyás, bolygatott talajon, korhadó növényi maradványokon, komposzton található meg. A lebontó szervezetek könnyen kiszorítják, és ezért nem kap életteret. Viszont ha valamilyen környezeti hatás folytán ezek a szervezetek viszszaoszorulnak vagy elpusztulnak, és életképes micélium kerül az adott táptalajba és a környezeti feltételek is adottak, akkor elindul a micélium-fejlődés benne. ...
<http://www.georgikon.pate.hu/lib/diploma.htm> (Origo-Vizsla)
12. Horváth Balázs: A CO² koncentráció és a hőmérséklet termésmennyiségre gyakorolt hatásai a lappangási szakaszban, a csiperketermesztésben. Szakdolgozat. Veszprémi Egyetem, Georgicon Mezőgazdaságtudományi Kar Növénytan és Növényélettani Tanszék, Keszthely 2002.
 ... A túl korán megkezdett szellőztetés hatására kis gombák mélyen képződnek, és szinte feltűrik a takaróanyagot. A borsónagyság elérésekor a szellőztetés mértéke az addigi kb. harmadára csökkenthető. Ebben a szakaszban adódik utoljára lehetőség a takaróanyag megöntözésére, mivel a szedés előtti napokban kiadott víz hatására a gomba megnyúlik, és érzékenyebbé válik a mechanikai hatásokra, így szedés közben könnyen elszíneződik. ...
<http://www.georgikon.pate.hu/lib/diploma.htm> (Origo-Vizsla)
13. Gömöri Zsolt: Pótlólagosan adagolt dúsítóanyag szerepe a csiperketermesztésben. Szakdolgozat. Veszprémi Egyetem, Georgicon Mezőgazdaságtudományi Kar Növénytan és Növényélettani Tanszék, Keszthely 2003.
 ... A komposzthoz adagolt dúsítót, vagy a csírázással egy menetben, vagy a csírázástól számított 10-12. napon adagolják. Általában 0,3%-os formalinnal kezelt, granulált szójaliszttel alkalmaznak 1 kg/100kg mennyiségben. Ilyen dúsítóanyag például a Millichamp 6000, amelynek külföldi és magyar kutatóintézetek vizsgálata alapján – mint a Horsti Gombatermesztési Kutatóintézet, vagy a Budapesti Növényegészségügyi Állomás – terménővelő hatása van. ...
<http://www.georgikon.pate.hu/lib/diploma.htm> (Origo-Vizsla)
14. Gombatermesztés az ókori Európában
 ... A termesztéssel foglalatostkodók némelyike valószínűleg felfigyelt a "termeszthető csiperke" tulajdonságára: emberi segédlet nélkül megkapaszkodó képességére a lótrágyán. Athénaiosz a Nikandroszt követő „gondolkodó” embereket

- említ. Közöttük minden bizonnyal akadtak olyanok, akik például a saját trágyagyűjtőhelyükön a spontán megjelent csiperkék gombafonalaktól „penészes” lótrágyáját fáik lombosára alá telepítették át. Lehettek, akik az utak mellett talált csiperketelepekből kapartak vagy ástak ki részeket, és vitték azokat haza. ...
<http://www.hhrf.org/uh/2002/uh13079.htm> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
15. Geml József: Molekuláris filogenetikai vizsgálatok és termesztési kísérletek vadon termő csiperke (*Agaricus*) taxonokkal. Doktori disszertáció. Budapesti Corvinus Egyetem Kertészettudományi Doktori Iskolája, Budapest 2005.
 ... A kétspórás csiperke életciklusát széleskörűen tanulmányozták már az 1970-es években is (RAPER et al. 1972, ELLIOTT 1972). A legtöbb csiperkénél, mint a bazídiomos gombák nagy részénél általában, a termőtest bazídiumain képződő négy bazidiospórába, a kariogámia és a meiózis után, egy-egy sejtmag kerül, így ezekből a spórákból kihajtó homokarionok csak az egyik párosodási típus genetikai hátterét hordozzák.
 ...
www.lib.uni-corvinus.hu/phd/geml_jozsef.pdf (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
16. Pallas Nagy Lexikona
 ... Csiperke vagy cseperke, csöpörke-gomba, pecserke, csiperke-galóca (*Agaricus campestris* L., vagy *Psalliota c. Pers.*; franc. champignon, l. a gombák képén), valamennyi ehető, kalapos azaz igazi gomba közt a leghasználatosabb. Földünk északi mérsékelt égövének nyirkos legelőin, kertekben, réten, erdőben és megtrágyázott gyepen májustól októberig igen közönséges. Tönkje a kalap középpontjával függ össze, 1,3-5 cm. hosszú, 0,6-2,5 cm. vastag, tömött. Spórája a kalap alsó színén levő, sugaras helyzetű lemezekén képződik. ...
<http://mek.oszk.hu/00000/00060/html/023/pc002380.html> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
17. Vízhányó Tünde: Colour Analysis for Mushroom Disease Assessment. Doktori értekezés összefoglalója. Szent István Egyetem Élelmiszertudományok Doktori Iskolája, Gödöllő 2000.
 ... A csiperkegombát számos mikróbas támadás érheti, melyek hatására a gomba felületén a betegség barnás foltok, illetve szöveti elváltozások formájában jelentkeznek és teszi a gombát értékesítésre alkalmatlanná. A kétféle eredetű barnulás (természetes és betegség okozta) megkülönböztetése volt a doktori munkám feladata. Az elkülönítéshez csak a színinformációt használtam, alapul véve a barnulások spektrális jellemzőit, illetve a képfeldolgozás kínálta lehetőségeket. ...
<http://www.doktori.jegyzek.hu/jeg.php?subid=1&pid=383&back=3> (Góliát, Altavizsla)
18. A gombatermesztés fejlesztése a molekuláris biológia módszereivel. Szerződésszám: OMFB-00352/02 (BIO-00046/01)
 ... Vadon termő *Agaricus bisporus* törzseket gyűjtünk be hazai, ill. külföldi génbankokból és hazai természeti környezetből. Ezekből szövettenyészeteket, ill. szaporítóanyagokat készítünk termesztési, nemesítési és molekuláris genetikai kísérletek céljából. Többek között az USA-ban (*Agaricus Resource Program*) és Franciaországban (INRA) begyűjtött törzsek igen értékes tulajdonságokat hordoznak. Magyarországon is célszerű egy ilyen génbank létrehozása a már meglévő (gén-)gyűjtemény kibővítésével. ...
<http://www.nkth.gov.hu/letolt/k+f/fitotechnologia2001.doc> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)
19. Szabó G., Rajkó R., Hodur C.: Csiperkegomba konvekciós és mikrohullámú szárításának összehasonlító vizsgálata. Szegedi Tudományegyetem, Szegedi Élelmiszeripari Főiskolai Kar Élelmiszeripari Műveletek és Környezettechnika Tanszék. Konferencia előadás. 4. Magyar Szárítási Szimpózium, Mosonmagyaróvár, 2001. október 18-19.
 ... Kutatásunk célkitűzése volt olyan kéméletes eljárás kidolgozása, amelynek eredménye új típusú élvezeti (gyorsan oldódó - instant és/vagy könnyű - snack-szerű) tulajdonságokkal rendelkező termék előállítására, ún. váltakozó rendszerű, konvektív és mikrohullámú szárítás kombinálását magába foglaló technológiával. Kritériumként tekintettük, hogy a termék a vízszanedvesítés után az eredeti friss gombához – annak fogyaszthatósági tulajdonságait tekintve – a lehető legteljesebb mértékben hasonlítson. ...

<http://www.szeff.u-szeged.hu/~rajko/pdfs/SzaboRajkoHodur4MSzSz.pdf> (Releváns hiperhivatkozás, amit egyik magyar nyelvű kereső sem keresett vissza.)

amit egyik magyar nyelvű kereső sem keresett vissza.)

21. Viczián Zsolt: Kertészeti-élelmiszeripari minőségbiztosítási rendszerek informatikai támogatása. Doktori értekezés összefoglalója. Budapesti Műszaki és Gazdaságtudományi Egyetem, Budapest 2000.

... A gombanövekedés folyamatának nyomon követésére olyan mérés technikát határoztam meg, melynek kivitelezése olcsó, ugyanakkor a növekedési folyamatra gyakorolt hatása elhanyagolható. A mérés technikát, mérés kiértékelést támogató szoftverek megtervezésével és elkészítésével egészítettem ki. A készített szoftverek alkalmasak a zsákon termesztett csiperkegomba felülnézeti fényképei alapján a kalapátmérő eloszlás automatikus, illetve félautomatikus meghatározására. ...

<http://www.chemonet.hu/hun/food/phd/bme/viczian.html> (Góliát, Altavizsla, Kurzor)

20. Zárójelentés. OTKA nyilvántartási szám: F 030634.

... A kutatás elsődleges célja módszer kidolgozása volt csiperkegomba (A. bisporus) látható minőségi jellemzőinek számítógépes látórendszerrel történő mérésére, a módszert alkalmazva a jellemzők mérésére az egyes minőségi osztályokkal összerendelve. A feltárt összefüggések alapján a csiperkegomba minőségének becslése. A számítógépes képfeldolgozás segítségével automata ellenőrző robotok építhetőek, amelyek minimális felügyelettel és beavatkozással képesek a betegségeket korán azonosítani, valamint az automatizált betakarítás során a minőségi osztályokba sorolást elvégezni. ...

www.uni-corvinus.hu/~blaszlo/data/pdf/kutatas_otka_f030634.pdf (Releváns hiperhivatkozás,

2. függelék

Az egyes keresőkulcsokra visszakeresett releváns és összes találatok száma

Rákóczi-szabadságharc	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Heuréka	5	3	2	836
Origo-Vizsla	7	3	5	4 240
Kurzor	11	3	2	157
Góliát	6	0	1	16 641
Altavizsla	6	1	1	60 894

1. táblázat.

„Rákóczi-szabadságharc” keresőkulcsra kapott találatok száma

Rákóczi-szabadságharc	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Heuréka	2	2	3	11 247
Ok.hu	14	9	9	7 490

Rákóczi-szabadságharc	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Kurzor	4	1	0	8 196
Góliát	5	1	1	16 641
Altavizsla	3	1	1	60 894

2. táblázat.

„Rákóczi-szabadságharc” keresőkulcsra kapott találatok száma a kontroll-teszt során

Évészavar	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Heuréka	6	1	0	125
Origo-Vizsla	6	5	6	1 070
Kurzor	9	2	0	1 422
Góliát	6	1	0	5 296
Altavizsla	6	1	0	5 296

3. táblázat.

„Évészavar” keresőkulcsra kapott találatok száma

Szököár	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Heuréka	3	1	4	446
Origo-Vizsla	7	4	4	6 560
Kurzor	5	1	3	12 766
Góliát	2	1	0	247
Altavizsla	2	1	0	247

4. táblázat.

„Szököár” keresőkulcsra kapott találatok száma

Kegyhely	Releváns találatok száma az első 20 találatban	Releváns találatok száma a második 20 találatban	Releváns találatok száma a harmadik 20 találatban	Összes találat száma
Heuréka	6	3	6	255
Origo-Vizsla	8	6	2	2 790
Kurzor	8	3	2	3 469
Góliát	7	4	2	5 286
Altavizsla	7	4	2	5 286

5. táblázat.

„Kegyhely” keresőkulcsra kapott találatok száma