

# Hogyan mérhető az internetes keresők szolgáltatása?

TÓTH ERZSÉBET

Nyíregyházi Főiskola, Matematika és Informatika Intézet  
tothe@nyf.hu

**Kulcsszavak:** internetes keresők, teljesítmény értékelés, minőségbiztosítás

A megjelenő értékelésekben a kutatók az internetes keresők minőségét számos mérőszám alapján mérik. Írásomban a keresők mérését a lehetséges mérési attribútumok irányából közelítem meg, hiszen azok együttesen határozzák meg a mérendő minőséget. Ezeknek a megfelelő mérőszámoknak a megtalálása rendkívül nehéz feladat, számos vita folyik erről a kutatók körében. Két olyan vizsgálat eredményeiről szeretnék beszámolni, amelyek újszerű mérőszámok bevezetését szorgalmazzák a mérésekbe. Az egyik kísérlet a SERVQUAL minőségbiztosítási modellt alkalmazza a keresők szolgáltatás minőségének a meghatározására. A másik vizsgálat teljes mértékben automatizálja a keresők mérését objektív mérési attribútumok használatával.

## 1. Bevezetés

Az internetes keresőkkel kapcsolatos értékelésekben a kutatók a keresők minőségét számos mérőszám alapján mérik. Írásomban az internetes keresők mérését a lehetséges mérőszámok irányából közelítem meg, hiszen azok együttesen határozzák meg a mérendő minőséget.

Két olyan vizsgálat eredményeiről szeretnék beszámolni, amelyek a web sajátosságaihoz igazodó, újszerű mérőszámok bevezetését szorgalmazzák a mérésekben. Az egyik kísérlet a SERVQUAL minőségbiztosítási modellt alkalmazza az internetes keresők szolgáltatás minőségének meghatározására. A másik vizsgálat teljes mértékben automatizálja a keresők mérését objektív mérőszámok használatával. Ezeknek a megbízható mérőszámoknak a megtalálása rendkívül nehéz feladat, számos vita folyik erről a kutatók körében.

Oppenheim [11] és Froehlich, T. J. [6] egy olyan szabvány kidolgozását javasolják, amely az alábbi mérőszámokat tartalmazná:

1. pontosság;
2. relatív teljesség;
3. válaszidő, amelyet naponta többször mérünk és rögzítjük a tesztelésre fordított időt is: a kereső milyen könnyen érhető el?;
4. egy adott időtartamon belül a találatok megbízhatóságának mérése: vajon mindig ugyanazokat a találatokat kapjuk-e az adott keresőkérdésre;
5. a zsákutcás, halott hivatkozások aránya;
6. az ismétlődő találatok aránya;
7. találatok minősítése felhasználókkal;
8. a grafikus felhasználói felület értékelése (annak használata mennyire felhasználóbarát?);
9. a súgó és a keresőprogram mennyire hasznos a kezdő és a tapasztalt felhasználók számára;
10. találatmegjelenítés;

11. reklámok jelenléte;
12. téma lefedettség;
13. a várható keresési lépések száma;
14. a találati leírások terjedelme és olvashatósága.

Egyéb értékelési szempontok:

15. Testreszabhatóság: a keresőszolgáltatás testreszabható-e?
16. Vizuális egyértelműség: a keresőkérdés és a találatmegjelenítés világos, érthető és következetes-e?
17. Navigáció: egyértelmű a navigációs eszközök használata? Könnyen tudunk-e a találatmegjelenítéstől a forrásdokumentumig eljutni és fordítva?
18. A keresőszolgáltatás felkínál-e valamilyen tematikus webhelylistát, amiben böngészhetünk?

Ez a lista természetesen bővíthet további új mérési szempontokkal. A mérőszámok általános szabványának hiánya nagy problémát jelent az értékelésekben. Ezen hiányosság miatt a keresőszolgáltatások értékelésével foglalkozó kutatás jelenleg nem egységes az alkalmazott módszerek tekintetében, ezért nagy szükség lenne megfelelő szabvány kidolgozására és használatára.

## 2. Az internetes keresők értékelésénél felmerülő problémák

Az internetes keresők elemzésére irányuló kísérletek nem egységes módszerekkel valósultak meg, ezért a legtöbb kutatási beszámoló tájékoztató jellegűnek tekinthető. Eddig főként a robotokkal felszerelt keresőkön végeztek ilyen jellegű méréseket, de elvileg bármilyen típusú kereső értékelhető. Leighton és Srivastava [10] szerint sok összehasonlító értékelés ellentmondásos következtetésekre jutott arra vonatkozóan, hogy melyik szolgáltatás nyújtja a legrelevánsabb találatokat.

Ezenkívül több tanulmány kisebb tesztelési kísérletek eredményeire épült, és nem számolt be az általa használt módszertanról [11]. Mivel ezek a tanulmányok kevés tesztelési eredménnyel rendelkeznek, ezért azok nem alkalmasak a mélyebb szintű statisztikai elemzések elvégzésére.

Su [17] megállapítja, hogy hiányzik a szisztematikus megközelítés ezekből a tanulmányokból és rámutat arra, hogy a kutatóknak nincs egységes kialakult véleménye arról, hogy mit mérjenek, és hogyan mérjék a szolgáltatást. Megjegyzi, hogy a legtöbb tanulmányból a felhasználók, mint aktív közreműködők kimaradnak. Általában az első 10 vagy 20 lekérdezett találat relevanciáját mérik. A relevancia ítéleteket többnyire a kutatók hozzák meg, nem pedig a felhasználók. A mérési eredmények rendszerint jelzik azt, hogy a legjobbnak minősített második vagy harmadik kereső között a teljesítménybeli különbség minimális.

A weblapok relevanciájának értékelésénél a legnagyobb problémát az eredmények befolyásolásának a kivédése jelenti. Előnyös megoldásnak tűnik, hogy ha az értékelő nem tudja, hogy a kapott találat melyik keresőszolgáltatástól származik. Ez a módszer csupán a weboldalak kezdeti ellenőrzésekor működik jól, hiszen a későbbi ellenőrzés és a frissítés már a forrásoldal ismerete mellett történik. A keresési folyamat sokféleképpen befolyásolható, például választható egy olyan tárgykör, amelyről köztudott, hogy az egyik kereső színvonalasabb szolgáltatást nyújt róla, mint a másik [10].

Fontos, hogy az értékelők ne legyenek elfogultak egyik keresővel szemben sem, és megőrizték pártatlanságukat objektív módszerekkel. Ennek a szemléletnek tükröződnie kell a keresőkérdések megválasztásában is. Ez a kritika természetesen egyaránt vonatkozik a hagyományos információ-visszakereső rendszerekben végzett tesztelésekre is.

Az értékeléseknél külön gondot jelent az internetes keresők állandó változása, mivel azok gyakran fejlesztik a keresési mechanizmusait és a felhasználói felületüket. Ehhez párosul még az a tény, hogy a világháló egy dinamikusan változó közeg. Ennek eredményeként a megjelenő értékelések rendkívül rövid életűek, és pillanatnyi helyzetképet tükröznek a keresőszolgáltatásokról. Mindezek ellenére igenis van értelme elemezni a használatban lévő keresőket, de szabványos értékelési módszereket erre a feladatra nem alkalmazhatunk. Általában megfigyelhető az a jelenség, hogy az elemzések egyéni értékelési módszerekről számolnak be, és többnyire elkerülik a szabványos értékelési módszerek használatát.

Az információ-visszakereső rendszerek hatékonyságának és teljesítményének mérése a cranfieldi modell alapján történik, ami a relevancia két fontos mérési szempontjára épül: a teljességre és a pontosságra. Mindkét mérőszám érzékeny arra, hogy a relevanciát hogyan definiáljuk és mérjük. A teljesség mérése megköveteli, hogy az értékelők a kapott találatok teljes halmazához hozzáférhessenek, vagy a találatokat képviselő mintához.

Ennek a követelménynek a kielégítése külön problémát jelent az internetes keresők értékelésénél [11]. Sokan érveltek amellett, hogy nem lehet mérni a teljességet, mert nehéz meghatározni a releváns találatok összességét egy adott lekérdezésre a weben. Az eddig megjelent tanulmányok csak a pontosságot vették figyelembe. Minimális mértékben, vagy egyáltalán nem kísérelték meg a különböző keresők teljességének mérését [3].

A pontosság mérése is nagy gondot jelent, mert az főként emberi relevancia ítéletektől függ, ezért az rendkívül szubjektív [1]. Azonban ennek a mérése egyértelmű, hiszen a keresési találatok megvizsgálása után azokat a releváns, illetve a nem releváns dokumentumok halmazába soroljuk [11].

Spink és Greisdorf [16] szerint jobb eredményeket nyerhetünk azáltal, hogy ha a relevancia többféle szintjét definiáljuk. Ezeket a relevancia szinteket már korábban is felhasználták az internetes keresők pontosságának a méréséhez [3, 5, 9].

### 3. A keresők mérésében használt új mérőszámok

Bar Ilan egy olyan mérés tapasztalatairól számolt be, ahol egyetlenegy lekérdezést futtattak le 6 keresőn (AltaVista, Excite, Fast, HotBot, Google, Northern Light) 33 alkalommal egy éven keresztül. Az értékelők egy jól körülhatárolt, stabil témát választottak lekérdezésre. Úgy fogalmazták meg a lekérdezést, hogy azt mindegyik kereső azonosan értelmezze. Ezzel is igyekeztek elkerülni a szótóképzéssel és a többszavas keresőkérdések kezelésével kapcsolatos problémákat. A mérés elsődleges célja az volt, hogy megvizsgálják a keresők működését hosszabb időtartamon keresztül. Találat-szolgáltatásuk stabilitását számos mérőszám irányából közelítették meg [1].

Úgy véljük, hogy ennek a mérésnek a legértékesebb részét képezi az itt alkalmazott, új mérőszámoknak a részletes tárgyalása. Ebben a részben bemutatjuk ezeket az újonnan bevezetett mérőszámokat a pontos meghatározásaikkal együtt.

#### 3.1. Technikai pontosság lekérdezésenként és a tanulmányozott időszak alatt

A relevancia megítélése helyett egy objektív mérőszámot alkalmazhatunk, amely megadott keresőszavakra keres a dokumentumban. Ez a mérőszám nem csupán objektív, hanem könnyen és gyorsan kiszámítható egy egyszerű minta megfelelő algoritmussal. Egy dokumentum, akkor minősül *technikailag relevánssá*, amennyiben az kielégíti a lekérdezés feltételeit. Vagyis azok a keresőszavak és kifejezések, amelyekről feltételezzük, hogy jelen vannak a dokumentumban, azok ténylegesen előfordulnak abban. Mindazok a keresőszavak és kifejezések, amelyekről feltételezzük, hogy hiányoznak a dokumentumból, azok nem is jelennek meg abban.

Egy URL-t, akkor tekintünk technikailag relevánsnak, amennyiben az technikailag releváns dokumentumot is tartalmaz. A *technikai pontosságot* úgy határozhatjuk meg, hogy a lekérdezett dokumentumok teljes számából kiszámítjuk a technikailag releváns dokumentumok százalék arányát. A technikai relevancia előnye az, hogy automatikusan kiszámolható, és nagy dokumentum halmazokra alkalmazható. Egyaránt alkalmas a „rossz dokumentumok”, azaz a nem létező URL-ek, és a lekérdezést ki nem elégítő dokumentumok kiszűrésére. Hátránya az, hogy nem állapítja meg a dokumentum fontosságát és megbízhatóságát. A fenti meghatározás egy URL technikai relevanciájára vonatkozik egy adott lekérdezésnél.

Ez a meghatározás sokféleképpen kiterjeszthető a vizsgált időszakra: egy URL, akkor tekinthető technikailag relevánsnak a vizsgált időszakban, hogy ha az annak minősült az első lekérdezéskor is. Másik lehetséges megoldás, hogy ha összegyűjtjük a technikailag releváns URL-eket minden egyes lekérdezéskor, vagy azokat az URL-eket, amelyek legalább egy lekérdezéskor minősültek technikailag relevánsnak. Bármelyik megoldás választható.

### 3.2. Relatív téma lefedettség lekérdezésként és a vizsgált időszak alatt

Egy kereső *relatív téma lefedettségét lekérdezésként* úgy számíthatjuk ki, hogy az adott lekérdezés során az egy adott kereső által megtalált technikailag releváns URL-ek számát elosztjuk az adott lekérdezés során az összes vizsgált kereső által megtalált technikailag releváns URL-ek számával.

Egy kereső *teljes relatív téma lefedettségét* úgy kapjuk meg, hogy az összes lekérdezés során az egy adott kereső által megtalált technikailag releváns URL-ek számát elosztjuk az összes lekérdezés során az összes vizsgált kereső által megtalált technikailag releváns URL-ek számával.

Ez a két mérőszám durván becsli a keresőszolgáltatás téma lefedettségét, azaz, hogy az mennyire képes a lekérdezésnek megfelelő találatokat szolgáltatni. A relatív téma lefedettség becsli a kereső téma lefedettségét lekérdezésként egy adott időpontban. Míg a teljes relatív téma lefedettség megállapítja a kereső relatív téma lefedettségét az egész időtartam alatt. A két mérőszám nagymértékben függ a vizsgált keresőeszközök csoportjától.

### 3.3. Új és teljesen új URL-ek

Egy URL akkor minősül *újnak* egy keresőszolgáltatás számára az adott lekérdezéskor, hogy ha azt a kereső egyetlenegy korábbi lekérdezésnél sem kérdezte le. Ezt a mérőszámot csak a második lekérdezéstől kezdődően lehet kiszámolni, azaz csak az utolsó (n-1) lekérdezésben kapott URL-eket kell figyelembe vennünk. Tehát az első lekérdezés mindig alapul szolgál, amihez viszonyítjuk a következő lekérdezést. Az új URL-ek mérőszáma azt jelzi számunkra, hogy a kereső milyen mértékben bővíti adatbázisát a témát illetően.

$új\ URL-ek_{(q,i)} = |\{az\ i.\ keresési\ alkalom\ során\ kapott\ technikailag\ releváns\ URL-ek\ száma\} - \{a\ j.\ keresési\ alkalom\ során\ kapott\ technikailag\ releváns\ URL-ek\ száma,\ ahol\ j < i\}|$

Az új URL-ek halmaza további két csoportra osztható: a *teljesen új URL-ekre* és az *újonnan felfedezett URL-ekre*. Egy URL, akkor tekinthető teljesen újnak egy kereső számára az adott lekérdezésnél, amennyiben azt korábban egyetlenegy kereső sem kérdezte le. A teljesen új URL-ek halmazát olyan URL-ek halmazának is tekinthetjük, amelyeket először kérdeztek le egy adott lekérdezésnél.

$teljesen\ új\ URL-ek_{(q,i,s)} = |\{az\ i.\ keresési\ alkalom\ során\ az\ s\ kereső\ által\ lekérdezett\ technikailag\ releváns\ URL-ek\ száma\} - \{az\ i.\ keresési\ alkalom\ előtt\ lekérdezett\ URL-ek\ száma\}|$

$újonnan\ felfedezett\ URL-ek_{(q,i,s)} = új\ URL-ek_{(q,i)} - teljesen\ új\ URL-ek_{(q,i,s)}$

Újonnan felfedezett URL-eknek minősülnek azok a korábban létező és régóta releváns weboldalak, amelyeket a kereső csak mostanában fedezett fel és adott hozzá a saját adatbázisához. Ezek a felsorolt mérőszámok a téma gyarapodását tükrözik a világhálón.

### 3.4. Elfelejtett, újra megtalált, elveszített, jól kezelt és rosszul kezelt URL-ek

A keresők sokszor tévesen távolítják el találatlistájukból azokat az URL-eket, amelyek továbbra is léteznek a weben és technikailag relevánsak. Az ilyen URL-eket *elfelejtett URL-eknek* nevezzük. Az elfelejtett URL-ek keresőnként és lekérdezésként egyaránt meghatározhatók. Azok azonban definiálhatók lekérdezésként, a keresőtől függetlenül is.

$elfelejtett\ URL-ek_{(q,i)} = |\{(i-1)\ keresési\ alkalom\ során\ kapott\ technikailag\ releváns\ URL-ek,\ amelyek\ továbbra\ is\ léteznek\ a\ weben\ és\ technikailag\ relevánsak\ az\ i.\ keresési\ alkalomnál,\ de\ azok\ mégsem\ kerülnek\ lekérdezésre\ az\ i.\ keresési\ alkalom\ során,\ ahol\ i > 1\}|$

Az elfelejtett URL-ek később újra felfedezhetők a következő lekérdezésnél. Egy URL csakis a legelső alkalommal számít újra megtaláltnak, miután az tévesen kimaradt a találatlistából. Ezeket az URL-eket *újra megtalált URL-eknek* nevezzük. Az újra megtalált URL-ek keresőnként és lekérdezésként egyaránt meghatározhatók. Azok azonban definiálhatók lekérdezésként, a keresőtől függetlenül is.

$újra\ megtalált\ URL-ek_{(q,i)} = |\{a\ j.\ keresési\ alkalom\ során\ kapott\ technikailag\ releváns\ URL-ek,\ amelyek\ tévesen\ maradtak\ ki\ az\ i.\ keresési\ alkalom\ során,\ ahol\ i < j\ és\ azok\ a\ technikailag\ releváns\ URL-ek,\ amelyek\ nem\ kerültek\ lekérdezésre\ a\ (j-1)\ keresési\ alkalom\ során\}|$

Azon URL-ek halmazát, amire semmilyen tartalmi megfeleltetést nem találunk az *elveszített URL-ek* halmazának nevezzük. Az információ teljesen elveszik ezekben az URL-ekben a felhasználó számára az adott

lekérdezésnél. Az elveszített URL-ek keresőnként definiálhatók, valamint kiterjeszthetők a lekérdezésekre is. Egy URL többször is lehet elfelejtett vagy újra megtalált a vizsgált időszak során.

*elveszített URL-ek* $_{(q,i)}$  =  $\{ \text{az } i. \text{ keresési alkalom során tévesen eltávolított URL-ek, amelyeknek nincs más tartalmi megfeleltetésük a } q \text{ lekérdezésre} \}$

A *jól kezelt URL-ek* halmazába tartoznak azok az URL-ek, amelyeket a kereső folyamatosan megtalál az első lekérdezés óta, vagy azok az URL-ek, amelyeket a kereső nem őríz meg, mert azok nem léteznek a weben, illetve technikailag nem relevánsak többé.

*jól kezelt URL-ek* $_{(q)}$  =  $\{ \text{azok a } q \text{ lekérdezésre kapott technikailag releváns URL-ek, amelyeket tévesen soha nem távolítottak el a keresési időszak alatt} \}$

A *rosszul kezelt URL-ek* halmazába tartoznak azok az URL-ek, amelyeket legalább egyszer tévesen elhagytak a találatlistából a keresési időszak alatt. A rosszul kezelt URL-ek halmaza két további halmazra osztható: a *rosszul kezelt, újra megjelenő URL-ek* halmazára (ez megegyezik az újra megtalált URL-ek halmazával, de annál bővebb halmazzal van szó) és a *rosszul kezelt, eltűnt URL-ek* halmazára (ez megfelel az elfelejtett, de újra meg nem talált URL-ek halmazának, de annál nagyobb halmazzal van szó). A rosszul kezelt URL-eknél csak az első ( $n-1$ ) lekérdezésben talált URL-eket vesszük figyelembe, mert egy URL-t nem tekinthetünk rosszul kezeltnek, amennyiben azt az utolsó lekérdezésben először fedeztük fel.

*rosszul kezelt URL-ek* $_{(q)}$  =  $\{ \text{U tévesen eltávolított URL-ek, ahol } i > 1 \}$

A Bar-Ilan által vezetett mérésben az alábbi módszerrel ellenőrizték, hogy a keresők helyesen távolították-e el az URL-eket a találatalmazukból. Mindegyik keresőt speciálisan kezelték a lekérdezéseknél és a kapott találatokat elmentették. A lekérdezett találatokról leválasztották az URL-eket Visual Basic program segítségével. Excelben lefutattak egy Visual Basic modult, ezáltal létrehoztak egy egyedi URL listát minden egyes lekérdezésre vonatkozóan. A hivatkozásokat lekérdezték és a találatok szövegét lementették a helyi merevlemezre egy „url2file” nevű segédprogrammal. Az aktuális lekérdezés és az azt megelőző lekérdezés eredményeit összehasonlították. Ezt követően megpróbálták felkutatni a hiányzó URL-eket az aktuális találatalmazban.

### 3.5. Az adatbázis változékonysága

Az *adatbázis változékonyságát* úgy számolhatjuk ki, hogy az  $i$ . keresési alkalom során lekérdezett technikailag releváns URL-ek számát elosztjuk a  $q$  lekérdezésre kapott technikailag releváns URL-ek összességével az egész keresési időszak alatt.

Ez a mérőszám a keresési találatok változékonyságát jelzi egy bizonyos időtartam alatt. Amennyiben annak magas a számértéke minden egyes keresésnél, az

azt jelzi számunkra, hogy a keresési találatok stabilak. Azonban annak túl magas számértéke arra int minket, hogy a kereső adatbázisa elavulttá válik. Ennek a mérőszámnak az optimális értékei se nem túl magasak, se nem túl alacsonyok.

*adatbázis változékonysága* $_{(q,i)}$  =  $\{ \text{az } i. \text{ keresési alkalom során lekérdezett technikailag releváns URL-ek száma} \}$  osztva  $\{ \text{a } q \text{ lekérdezésre kapott technikailag releváns URL-ek összessége az egész keresési időszak alatt} \}$

### 3.6. Azon keresések száma, amelyben az adott URL-t lekérdezik

Ez a mérőszám a keresési találatok változékonyságához kapcsolódik. Tulajdonképpen azt jelzi, hogy egy adott URL-t pontosan hány keresésnél kérdezték le. Definiálható keresőnként, illetve a keresők egy csoportja számára is.

### 3.7. Állandósult URL-ek

*Állandósult URL-eknek* tekintjük azokat a technikailag releváns URL-eket, amelyeket az összes keresés során megkapunk. Egy URL, akkor minősül technikailag relevánsnak, amennyiben az az összes keresés során technikailag relevánsnak tekinthető. A fenti mérőszámok többsége kiszámítható az állandósult URL-ek halmazára nézve.

Ez a halmaz lehetővé teszi, hogy tanulmányozzuk a keresők változásait egy bizonyos időtartam alatt, még pedig úgy, hogy a keresők változásait elkülönítjük a világháló dinamikus változásaitól. Azonban a vizsgált halmaz túlságosan kisméretű lehet az elemzés számára [1,2].

## 4. A SERVQUAL modell alkalmazása a keresők értékelésére

Xie, Wang és Goh egy minőségbiztosítási modellt adaptált az internetes keresők minőségi vizsgálatára. A SERVQUAL modellt Parasuraman, Zeithaml és Berry [12-15] fejlesztette ki azzal a céllal, hogy mérni lehessen egy szolgáltatás minőségét. Kezdetben a szolgáltatás minőségének 10 dimenzióját javasolták, később az empirikus kutatások ezt lecsökkentették 5 dimenzióra. Egy tizennégy változóból álló, öt dimenziós eszközt alakítottak ki az internetes keresők minőségének mérésére, amelyet az 1. ábra szemléltet [18].

Kérdőíves felmérést végeztek a felhasználók körében, amely két részből állt. Az első rész általános információkat közölt a felhasználókról és a kedvelt keresőszolgáltatásokról. A második rész az internetes keresőkkel kapcsolatos felhasználói elvárásokat és a tényleges, észlelt állapotot mérte.

Egy szolgáltatás alacsony, illetve magas minőségének megítélése attól függ, hogy a felhasználók hogyan érzékelik a szolgáltatás aktuális teljesítményét elvárásai függvényében. Ezért a felhasználói elvárásokat és a felhasználói észlelést különbség értékekkel, azaz



'gap'-ekkel hasonlították össze. Ezt a módszert különbségelemzésnek, azaz 'gap analysis'-nek hívják [13].

A pozitív különbség érték azt jelzi, hogy a szolgáltatás minősége kielégíti, vagy túllépi a felhasználói elvárásokat. A negatív különbség érték azt jelzi, hogy a szolgáltatás minősége nem tesz eleget a felhasználói elvárásoknak. A SERVQUAL különbségértékek kicsik voltak, -1.44 és 0.66 között mozgott az értékük. Cronbachnak [4] az elvárási szintre és az észlelési szintre vonatkozó megbízhatósági becslése 0.76 és 0.88 volt külön-külön. Ebből adódóan a tizennégy változóból álló összetevő általános értékeket nyújt az elvárási szintre és az észlelési szintre vonatkozóan.

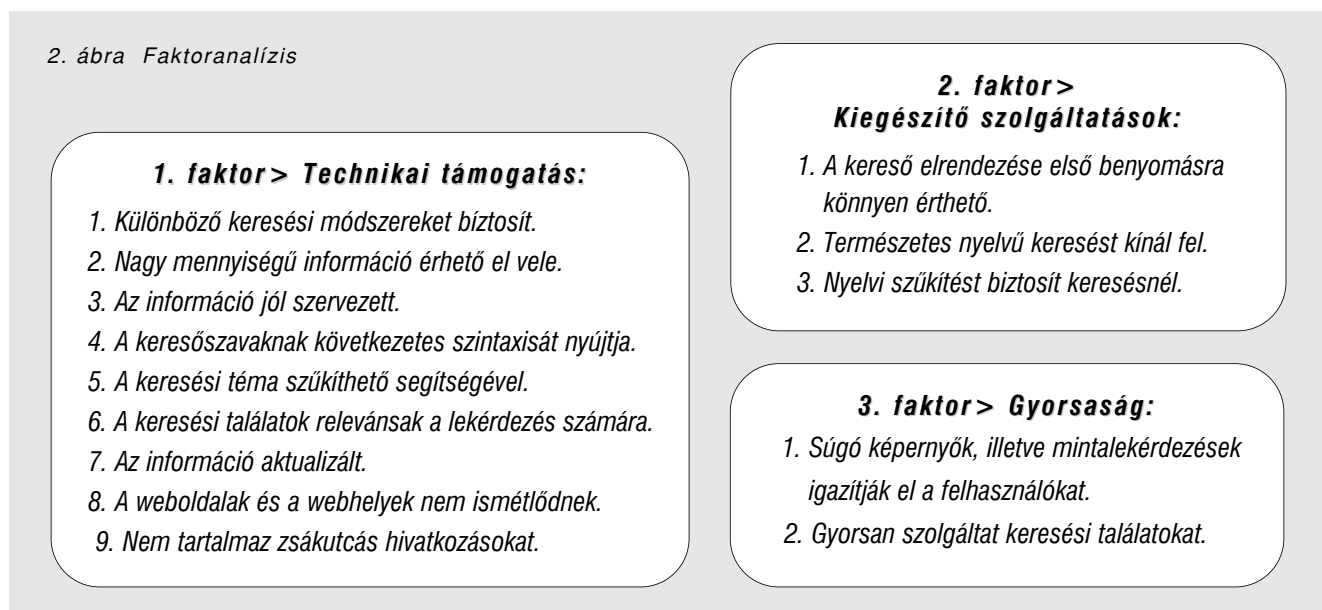
Az eredeti öt dimenziót mások nem alkalmazták egyélesen a megjelenő tanulmányokban. Ezért faktora-

nalízist használtak annak érdekében, hogy megállapítsák hány faktor, illetve főbb dimenzió fejezi ki legjobban a változók kezdeti halmazai között megfigyelt korrelációkat [7].

'Varimax' forgatásos módszerrel egyértelműen elválasztották a faktorokat egymástól és maximalizálták a betöltéshez szükséges szórásnégyzet összegeket [8]. Az öt dimenzió helyett összesen hármat azonosítottak be, amelyek faktoroknak minősültek. A három faktorhoz kapcsolódó 14 változót másképpen rendezték el, mint a korábban kifejlesztett öt dimenziós eszköz esetében. Ez a változás jól nyomon követhető a 2. ábrán.

A SERVQUAL modellben található *áttekinthetőséget*, *megbízhatóságot* és *garanciát* a felhasználók nem tekintették különálló faktornak. Forgatás után a betöl-

2. ábra Faktoranalízis



téshez szükséges szórásnégyzet összegek a *technikai támogatás* esetében 5.362-nek feleltek meg, ami a teljes variancia 38.3%-át tette ki. Ezért a technikai támogatást tekintjük a legnagyobb és a legfontosabb faktornak a három beazonosított faktor közül.

Három változót a felhasználók a *kiegészítő szolgáltatások* faktorhoz soroltak a korábbi *felhasználói támogatás* helyett. Ez az új faktor a teljes variancia 12.9%-ának felelt meg. A *gyorsaság faktor* csupán két változót tartalmazott és a teljes variancia 11.4%-át tette ki [19].

## 5. Összefoglalás

Az előzőekben ismertetett kutatási kísérlet rendkívüli jelentőségű, hiszen a minőségbiztosítás elvárásaitól indulva közelíti meg és szűri ki azokat a mérőszámokat, amelyek döntő fontossággal bírnak egy mérés megtervezésekor. A vizsgálat érdeme, hogy felhasználói észrevételekre támaszkodik a változók csoportosításánál és azoknak a megfelelő faktorhoz történő sorolásában.

A másik vizsgálat eredményei szintén elgondolkodtatóak, hiszen az automatizált technikák segítségével igyekszik minél hatékonyabban mérni a keresők teljesítményét, ezáltal pedig kiküszöböli a méréseknél felmerülő szubjektivitást. Minél több ilyen irányú kezdeményezésre lenne szükség a jövőben, hogy a mérőszámoknak egy általánosan alkalmazható szabványa jöjjön létre és ne egyéni kísérletezések történjenek ezen a területen.

### Irodalom

- [1] Bar-Ilan, J.: Methods for measuring search engine performance over time. In: Journal of the American Society for Information Science and Technology, Vol. 53., no.4., 2002, pp.308–319.
- [2] Bar-Ilan, J.: Criteria for evaluating information retrieval systems in highly dynamic environments. Paper at Web Dynamics Workshop, p.8.
- [3] Clarke, S. J.–Willett, P.: Estimating the recall performance of web search engines. In: Aslib Proceedings, Vol. 49., no.7., July/August 1997, pp.184–189.
- [4] Cronbach, L. J.: Coefficient alpha and the internal structure of tests. In: Psychometrika, 16 1951. pp. 257–334.
- [5] Ding, W.–Marchionini, G.: A comparative study of web search performance in global complexity: information, chaos and control. In: Proceedings of the 59th Annual Meeting of the American Society for Information Science, Baltimore, Maryland, 21–24. October 1996, pp.136–142.
- [6] Froehlich, T. J.: Case study 5.1: Developing search engine evaluation criteria. In: Library evaluation, Libraries Unlimited, 2001. pp.185–200.
- [7] Gorsuch, R. L.: Factor Analysis. Hillsdale: Lawrence Erlbaum Associates, 1983.
- [8] Hair, J. F. [et. al.]: Multivariate Data Analysis. Prentice-Hall, 1998.
- [9] Leighton, V. H.–Srivastava, J.: First 20 precision among World Wide Web search services (search engines). In: Journal of the American Society for Information Science and Technology, Vol. 50., no.10., 1999. pp.870–881.
- [10] Leighton, H. V.–Srivastava, J.: Precision among World Wide Web search services (search engines): Alta Vista, Excite, Hotbot, Infoseek, Lycos – 1997. www.winona.msus.edu/library/webind2/webind2.htm
- [11] Oppenheim, C.–Morris, A.–McKnight, C.–Lowley, S.: The evaluation of WWW search engines. In: Journal of Documentation, Vol. 56., no.2., March 2000, pp.190–211.
- [12] Parasuraman, A.–Zeithaml, V. A.–Berry, L.L.: A conceptual model of service quality and implications for future research. In: Journal of Marketing, Vol. 49. 1985., pp.41–50.
- [13] Parasuraman, A.–Zeithaml, V. A.–Berry, L.L.: Delivering quality service: balancing customer perceptions and expectations. New York, Free Press, 1990.
- [14] Parasuraman, A.–Zeithaml, V. A.–Berry, L.L.: Refinement and reassessment of the SERVQUAL scale. In: Journal of Retailing, Vol. 67., no.2., 1991, pp.420–450.
- [15] Parasuraman, A.–Zeithaml, V. A.–Berry, L.L.: SERVQUAL: a multiple item scale for measuring customer perceptions of service quality. In: Journal of Retailing, Vol. 64., 1988, pp.12–40.
- [16] Spink, A.–Greisdorf, H.: Regions and levels: Measuring and mapping users' relevance judgments. In: Journal of the American Society for Information Science and Technology, Vol. 52., no.2., 2001, pp.161–173.
- [17] Su, L. T.: Developing a comprehensive and systematic model of user evaluation of Web-based search engines. In: M. E: Williams (Ed.), National Online Meeting: Proceedings, 1997, pp.335–345., Medford, NJ: Information Today.
- [18] Xie, M.–Wang, H.–Goh, T. N.: Quality dimensions of Internet search engines. In: Journal of Information Science, Vol. 24., no.5., 1998, pp.87–94.
- [19] Wang, H.–Xie, M.–Goh, T. N.: Service quality of Internet search engines. In: Journal of Information Science, Vol. 25., no.6., 1999, pp.499–507.