



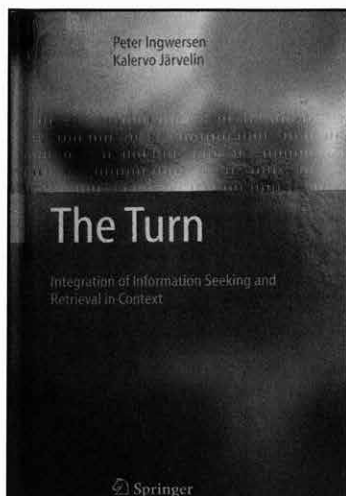
A fordulat

Információkeresés-kutatás és perspektívája:
a kognitív szemlélet

Jelentős összegző művet írt az információtudomány nemzetközi kutatói élvonalából ismert, hosszú évek óta együttműködő két szerző: *Peter Ingwersen* és *Kalervo Järvelin*. A dán- finn szerzőpáros hosszú kutatási periódus összegzésére, átértékelésére és átértelmezésére vállalkozott. Körültekintő és kritikus visszatekintéssel foglalják össze az *információ-visszakeresés* (Information Retrieval, IR)¹ és az *információkutakodás*, közkeletű szóval *információkeresés* (Information Seeking/Searching, IS) közel ötven éves kutatástörténetét. A szerzők a két terület (IR&S) eredményeit integráltan tárgyalják, sokrétűen elemzik az empirikus kísérletek eredményeit, hogy végül eljussanak az elméleti keretek körvonalazásához. A további fejlesztéseknél a kognitív szemlélet érvényesítését látják szükségesnek.

A kötet egyszerű – a „kemény” tudományok címadási szokásait némiképp fölragó irodalmias – főcímmel a szerzők a közel sem egyszerű témakörre szeretnének utalni. A címválasztáshoz az előszóban külön magyarázatot fűznek: „itt az ideje, hogy visszatekintsünk és előre is tekintsünk, hogy az információkeresésről és infor-

¹ A recenzióban az angol rövidítéseket használom.



INGWERSEN, Peter – JÄRVELIN, Kalervo

The turn : integration of information seeking and retrieval in context. – Dordrecht : Springer, 2005. – XIV, 448. p. – (The Kluwer international series on information retrieval ; 18)
ISBN 1-402-03850-X

1-4020-3851-8 ; 1-4020-3850-X ;
978-1-4020-3850-1 ; 978-1-4020-3851-8

e-kiadás: DOI: <http://www.springerlink.com/content/141827/>

máció-visszakeresésről új, integrált módon gondolkozzunk: a szűk, egymástól elszigetelt kutatási ösvényekről széles sugárútra kell rátérnünk.” (VII. old.) A fordulat a *kognitív fordulatvétel* szükségességét jelenti: a szerzők meggyőződése szerint a kognitív szemlélet kiterjesztése jelentheti a sok ponton megtorpanó és elavuló (könyvtári) hagyományokra alapozó információkereső rendszerek továbbfejlesztésének és korszerűsítésének zálogát.

A „kognitív fordulat” előzményei a 90-es évek elejére tehetőek: a recenziónk tárgyát képező *The Turn* Ingwersen 1992-ben publikált monográfiájának (*Information Retrieval Interaction*, London : Graham) folytatásaként, továbbfejlesztéseként is felfogható. A hangsúlyeltolódásokra a szerzők maguk hívják fel a figyelmet: míg az 1992-es monográfiában az interaktivitás, az egyén mint kereső személy kognitív struktúrája, valamint az interfész funkcionalitása mint *fontos* vizsgálandó jelenség jelenik meg, addig a 2005-ös műben ezek a kérdések az IR&S *fókuszba helyezett*, mélységeiben és komplexitásában bemutatott, dimenzionált kérdéseivé lépnek elő.

Az információtudományban a két információkereséssel foglalkozó terület más-más tudományterülettel ötvöződve, két elkülönülő ágazatban jelenik meg. Az *információ-visszakeresés* (IR) kérdéseivel a számítástudomány foglalkozik: ez az ágazat az információtechnológiára alapozva az információ-visszakereső eljárások, eszközök, algoritmusok és rendszerek fejlesztésére összpontosít. Tárgya a potenciális információ kinyerhetősége, reprezentációja, tárolása, a keresés, szűrés és megjelenítés technikája. Az *információkeresés* (IS) kutatói ezzel szemben tárgyakat általában a társadalomtudományokhoz sorolják, és könyvtártudományi megközelítéssel szemlélődnek. Ingwersen és Järvelin kísérletet tesz a két elszigetelődésre hajlamos kutatói közösség eredményeinek és szemléletmódjának összeolvasztására. Ezt a szándékot a kötet alcíme is jelöli: *az információkeresés és információ-visszakeresés integrációja összefüggésében*.

A szerzők meglátása szerint a két terület művelői nem sokat kommunikálnak egymással, és nem vesznek tudomást egymás módszereiről, sem eredményeiről, holott a hatékony keresőeszközök megteremtéséhez éppen annyira nélkülözhetetlen az emberi információs magatartás, a fel-

adatorientált kontextus figyelembevétele, mint amennyire megkerülhetetlen a problémamegoldó helyzet, a tudásállapotok, kognitív struktúrák vizsgálatánál azoknak a technikai eszközöknek a figyelembevétele, amelyeket az emberek információszerzésre használnak.

A szerzők komplex szemlélettel, az elméleti megfontolások és empirikus vizsgálatok sokrétű elemzésével és egymásra vetítésével kívánják segíteni a technikai eszközök, informatikai rendszerek és a felhasználói gyakorlat alakulását. A könyvben ezért kerül középpontba a kognitív szemlélet, az ember és gép közötti interakció, a problémamegoldó szituáció, a feladat, valamint kontextusai, a tudásszerzés módja, a tudástípusok megismerése, az információkeresés és az információ felhasználása, a modellek, keretrendszerek, ezek elméletei és kutatási módszerei.

A kiindulópont a laboratóriumi modellek alapján kifejlesztett visszakereső módszerek tesztelésének, kísérleti eredményeinek áttekintése és kritikus összegzése. A különféle algoritmusok hatékonyságának mérésére kifejlesztett laboratóriumi módszerek az ún. Cranfield-paradigmára alapozódnak. A paradigmára jellemző kísérleti módszerekkel az 1960-as évektől azt kutatták, hogyan lehet növelni a rendszerek keresési hatékonyságát. A laboratóriumi körülmények között végzett vizsgálatok jól körülhatárolható kísérleti tesztkollekcióval, adatbázison végezhetőek el. A visszakereső algoritmusok hatékonyságának fő értékelési szempontja a releváns dokumentumok, illetve a releváns információk visszanyerésének rendszer-képessége. A relevancia-értékek a kereső-algoritmus precizitásának és a visszanyerés teljességének mértékét fejezik ki.

A laboratóriumi IR vizsgálatok a számítástechnikai fejlesztések homlokterében állnak. Ezek a kutatások – amint azt a szerzők megállapítják – nem vesznek tudomást a humán információkeresés jellegzetességeit feltáró módszerekről és az IS-kutatások eredményeiről. A szerzők azt is belátják, hogy a laboratóriumi modellt sok kihívás és kritika érte, amiért a modell nem veszi figyelembe a valós szituációkat, a valós használókat, és nem tesz különbséget a magasabb rendű relevancia elemei között sem: (a) a kognitív, (b) a helyzeti, (c) az algoritmos és (d) a téma-relevancia között. A kereső személy feladatai, a feladatmegoldás fázisai, szituációs és kulturális környezete, valamint a kognitív adottságok nagyban befolyásolják a keresés menetét, stratégiáját, a használó taktikáját, így magának a relevanciának a helyzeti és más különféle szempontú értékét is.

A szerzők az *információ* fogalmának nem abszolút érvényű meghatározását keresik, hanem az információ megértését és használatát előtérbe helyezve alakítják ki az információ kognitív és kommunikációs szempontokat magában foglaló értelmezését. Leírásaikhoz bevezetik a *kondicionális kognitív információ* fogalmát. Fontosságot kap mind az előállított, mind a befogadott információ a környező aktorok, aktusok és kontextusok jellemzőivel, változóival együtt. Valamennyi aktossal mint aktív közreműködővel számolnak: az interakcióban a keresőrendszer is *befogad*, amikor a használó kérdését bináris jelekké transzformálja, feldolgozza és visszaküldi a bináris jelkészletből nyelvi jellé transzformált eredményhalmazt az értelmező embernek, az ember számára értelmezhető formában. A humán oldalon a kérdésfeltevő személy

a saját tudásszerkezetéből hoz létre információt, és ez alapján fogalmazza meg kérdését. A fázis ugyanúgy transzformációnak tekinthető, mint az a helyzet, amikor egy a rendszer által visszaküldött és megjelenített adatot a felhasználó megért, információvá alakít, befogad. A befogadó tudásállapotát az új információ meg is változtatja.

E folyamatok kapcsán tárgyalnak a szerzők olyan jelenségeket, mint a *kognitív szabadesés* (cognitive free fall) és *rekonstrukció* – a mondandónak egyik oldalról a kimondással/továbbítással járó kontextustalanodása, másik oldalról pedig a befogadó kogníciója során az üzenetkontextusának rekonstrukciója és interpretációja. A kontextus eltűnése, majd rekonstrukciója során az egyik aktor által előállított, továbbított majd a másik aktor által befogadott információ jelentés-tartalma jelentősen el is térhet egymástól. A jelenség illusztrálására a kötetben számos, a határterületek szakirodalmából jól ismert érdekes példát vonultatnak fel a szerzők.

A jelenséghez kapcsolható a könyvtári rendszer–használó nyelvének különbségéből eredő illeszkedési probléma. A használó ’tárgyszó’ mezőbe beírt kifejezése – amelyet a természetes nyelv szabályai szerint ad meg szabadszavasan, ragozott szókapcsolatokkal vagy akár teljes mondatokkal – a rendszer ’tárgyszó’ keresőpontjáról valójában a kontrollált, prekoordinált fogalmi szótárt (tezauruszt, formális ontológiát) hívja meg. Hasonló jelenség a használó próbaszerencse alapon végzett kulcsszavas kísérletére a válaszul kapott találatok értelmezhetetlen volta az emberi aktor előtt. Ember és gép nem egy nyelvet beszél: különbség van a szótár intellektuális előállítója (pl. egy osztályozó könyvtáros) és a kereső-aktor mentális ismeretei és nyelvhasználata között. Különbség lehet a szövegszavas

automatizált index, a relevancia-kalkulációval előállított lista, web-környezetben a horgony-információkat figyelő hálózati keresőgépek és a humán aktor szubjektív elképzelése között, mint ahogyan a különféle illesztő algoritmusok vagy a *pseudo-relevancia* visszacsatolási metódusokat² alkalmazó kérdésmódosítási eljárások is más-más eredményre vezethetnek. Érdekes ilyen szempontból az Ingwersen nevéhez kötődő relevancia-teljesítményt fokozó *polireprezentáció* elvén alapuló ún. *kognitív átfedés* (cognitive overlap) kiaknázása, melynek során a más-más kognitív eredetű és más-más funkciót szolgáló adatfélések megfeleltetésével érhető el magas relevanciaszint. Az eljárás például a dokumentum címstruktúrája, külső hivatkozásai (out-link), a dokumentumra mutató hivatkozások, vagy az azt idéző utalások (in-link), tezauruszok, folyóiratcímek, indexelők által hozzáadott deskriptorok, műfaji megjelölések stb. adatelemeinek tartalmi átfedéséből „keveri ki” a legrelevánsabb találatokat. A tudományos dokumentumokra vonatkoztatott modell szemléletes ábrája a könyv 207. oldalán található.

A holisztikus kognitív szemléletre való átváltás, vagyis a fordulat a két szerző szerint akkor veszi kezdetét, amikor feladjuk abbéli hitünket, hogy az információ-vissza-kereső eszközök a kognitív és funkcionális struktúrák változatosságát egységesen tudják majd kezelni. Ehelyett azt kellene tudomásul vennünk, hogy ezek a struktúrák inherensen különböznek, és mint ilyeneket kell őket tanulmányozni és kezelni. Sarka-

² *Pseudo-relevance feedback – a válaszok relevancia szerint rangsorolt listáján a legrelevánsabbnak tartott N számú találat (top-N), ld. a kötet 141. oldalán.*

latos megállapítás az is, hogy az ember–gép viszonyának aszimmetrikus kognitív természete akadályozza az információ-átvitelt, így különösen fontossá válik a használati felületek fejlesztése. Az interfésznek az interakció során segítenie kell az egyedi felhasználót, ugyanakkor a lépések, folyamatok követésével, a visszajelzések feldolgozásával, a rendszer tulajdonságainak folyamatos javítását is el kell érnie. A felületnek hármaskörű funkciója van: a) közvetít, b) biztosítja az intelligens visszakeresést, c) ha szükséges, tanítóként, tanácsadóként, segítőként kapcsolódik be a keresési folyamatba.

A szerzők sorra veszik az információs viselkedés (information behaviour), az információkeresésre irányuló kutatások eddig ismert modelljeit (többek között *T.D. Wilson, D. Ellis, C.C. Kuhlthau, T. Saračević, P. Vakkari—M. Kuokkanen, K. Byström—K. Järvelin* stb.). Valamennyi modellt kritikusan értékelnek, elemeznek, megvilágítva a modellekben azokat a pontokat, amelyek a javasolt kognitív szemléletben is hasznosíthatók vagy továbbfejleszthetők lehetnek. A kutatási meta-modellben a diskurzuselemzés is helyet kap, mellyel megérthető az aktorok tudásformációinak, háttérfeltevésének, magától értetődő interpretációs repertoárjainak szerepe a feladatvégzés közben, és megérthetők azok az emberi reakciók is, amelyeknek empirikus lenyomatait többek között a keresőfelületek naplófájlaiban is meg lehet figyelni.

A webkeresőgépek visszakeresési teljesítményének mérése a naplófájlok adatainak statisztikai feldolgozásával, a keresőszavak és a kérdések szintaxisának vizsgálatával, a navigációs irány követésével könnyűnek tűnik, ám ugyanúgy nehezen vizsgálható marad az emberi, gondolkodó, összefü-

géseket kereső értelmes lény mint kliens: a kereső előtudása, stratégiái, motivációi, problémamegoldó mechanizmusai és azok a környezeti tényezők, amelyek között a használati információt keres.

Az automatizmusokon alapuló empirikus adatgyűjtés, a statisztikai adatfeldolgozásból kinyerhető tipikus műveletsorok, használati keresési taktikák és stratégiák körvonalazása azonban önmagában még nem elegendő az ember problémamegoldásra irányuló információkereső tevékenységének megismeréséhez és modellezéséhez. A szerzők egy sor olyan laboratóriumi és más alternatív módszert, koncepciót, filozófiai és kognitív nézetet ismertetnek, amelyek elősegítik az ember–számítógép interakció megértését, a keresési–interpretációs folyamat összetevőinek tágabb látókörű vizsgálatát a cselekvési szituáció és tudásállapotok formálódásának szempontjából. Tovább bővítve a kört sorra veszik a szociális és kulturális kontextusok jellemzőit és változóit: az információszerzés szituációs természetét, a személyek alkalmazkodási képességét a felmerülő probléma megoldása közben, amelyet egyben a kulturális tájékozottság és a környezet is meghatároznak, miközben ezek a tényezők magukat az információs igényeket is meghatározzák. Mindezen tényezők összhatásának figyelembevételével dolgozzák ki a holisztikus kognitív szemléletű, kontextuális dimenziókba helyezett – és tegyük hozzá: manapság dominánssá váló – kutatási program keretrendszerét.

A könyv kilenc fejezete lépésről lépésre, egyre komplexebb köröket róva rajzolja ki e keretrendszer elméleti, kulturális, technikai-fejlesztési, vizsgálati-visszacsatolási és fejleszthetőségi összetevőit. A fejezetek, alfejezetek felépítése sematikus, a szerzők

a szöveg tartalmi tagolásának szerkezetével feltételezhetően a *Springer* kiadó támasztotta követelményeknek tesznek eleget, ugyanakkor ez a szerkezet a fejezeteket a moduláris feldolgozásra is alkalmassá teszi.

A fejezetek problémafelvetéssel kezdődnek, a kutatási eredmények összefoglalásával folytatódnak, kutatástörténeti és szakirodalmi áttekintést adnak; nem maradhat ki a kutatási módszerek, modellek, elért eredmények részletezése, elemző értékelése. A laboratóriumi modell tárgyalásmódját szem előtt tartva az a benyomásunk támadhat, hogy nagyobb hangsúllyal szerepelhetne a web IR&S témája és kritikus összefoglalása. A két szerző viszont inkább az elméleti összefüggéseket keresi, a modellezhető keretrendszerre fókuszál, alapelveket keres az IR&S kutatások modelljeinek tervezéséhez. Talán ezzel magyarázható, hogy a szerzők csak annyira érintik a divatossá váló webkeresés kutatásának témáit, amennyire azok felhasználhatók a szerzők támasztotta kritériumoknak megfelelő meta-modell megalkotásánál. Minden nagyobb téma lezárásánál a korlátok, nehézségek, erősségek, hátrányok és nyitott problémák, valamint a kortárs kutatási irányok is bemutatásra kerülnek.

A nyitó fejezet az IR laboratóriumi kutatásokat tekinti át. A második fejezet az információ fogalmát kognitív megközelítéssel értelmezi át, szem előtt tartva az IS&R alkalmazásokat. Elrugaszkodik a 80-as évektől domináló individualista–atomista felhasználó–központú nézetektől, amelyek a használót legfőljebb mint megfigyelőt veszik számba, attól rettegve, hogy a használói magatartás tanulmányozása csak káoszhoz vezetne. A könyv írói ezzel szemben a keresési folyamatban részt vevő

aktorok mindegyikére *közreműködőként* tekintenek. Figyelembe veszik az aktorok viselkedését és percepcióját befolyásoló kulturális közeget, megkülönböztetik az információszerzés különféle szintjeit is. A harmadik fejezet az IS-kutatásokat, a következő fejezet a rendszerközpontú, matematikai–logikai modellekre épülő IR-kutatásokat taglalja, részletesebb kitérőkkel a dokumentum–keresőkérdés–relevancia, indexelés–osztályozás–klaszterálás, felület–vizualizáció, relevancia-visszacsatolás–kérdésmódosítás–interakció, természetesnyelvű feldolgozás, szakértői rendszerek–interfészek és a relevancia-kiértékelés módszereire. Az ötödik fejezet a kognitív és használó–központú IR-rendszerekre irányítja a figyelmet. A webkörnyezet felé is kitekintve, részletezi a kognitív elmélet polireprezentációs alkalmazását, a keresők kognitív stílusait, a sztenderd és web IR interakciót és relevanciát.

A könyv lelke a hatodik fejezet, amelyben a szerzők eljutnak az IS és IR kutatást integráló modell bemutatásáig. A hetedik fejezetben a szerzők részletesen kifejtik az általuk javasolt kognitív keret alkalmazhatóságát az IS&R kutatásokban, mérlegelik a javaslat jelentőségét. Kilenc tág dimenziót tárgyalnak a problémamegoldó szituációtól az interfészekig. Végző konklúziójukban kifejtik, hogy az IR-kutatásoknak a számítógépes laboratóriumi kontrollkísérletek megtartása *mellett* kell figyelmüket a kontextuális és kognitív összetevők vizsgálatára fordítaniuk. A két utolsó, rövidebb fejezet további kutatási programokat, kutatás-csomagokat terjeszt elő, és tömören összefoglalja a végző konklúziókat.

Talán nem túlzás kimondanunk: a könyv *alapmű*. Jól használható mindazok számára, akik szakemberként szeretnék megis-

merni vagy mélyebben megérteni a gép és ember együttműködésén alapuló információkeresés mibenlétét, jelenlegi könyvtári rendszereink és az alternatív, más internetes eszközök tulajdonságait, képességeit, erősségeit és gyengeségeit, kutatási és fejlesztési távlatait. Kiterjedt bibliográfiát, gazdag, magyarázatokkal ellátott szemléletes ábranyagot, precízen definiált fogalomjegyzéket valamint tárgymutatót tartalmaz.

A *The Turn* teljes tartalma tanítható, részleteiben – akár különálló tantárgyi egységként – is, a felsőbb könyvtártudományi vagy más információtudományi szakok egyetemi és doktoriskolai évfolyamain, de egyszerűsítve, a bevezető tárgyak tananyagába építve is segítheti a naponta használt keresőeszközök lényeges jellemzőinek megértését. Színvonalas szakmai olvasmányként pedig gyakorló könyvtárosok,

informatikusok, modellezők, programozók, információstársadalom-kutatók és kommunikációs szakemberek számára is jó szívvel ajánlható.

Ingwersen és Järvelin sokoldalú, több mint két évtizedes munkáját összegző, a vonatkozó szakirodalmat alaposan integráló könyvének végén nincs pont. Természetesen, nem véletlenül becsúszott sajtóhibáról van szó: a könyv végén nincs pont olyan értelemben, hogy az az állomás, illetve a cím szerint sugallt fordulópont, ahová a szerzők elérnek, nem nyugvópont. Számos kutatási feladat, átgondolandó, nyitva maradt és megoldandó kérdés kínálkozik fel, amely az elmélyültebb érdeklődőt további tudományos kérdésfeltevésre vagy éppen séggel a kitérülő probléma kutatására is sarkallhatja.

Dudás Anikó



Megjelent az UTCA projekt ízelítőjének 2. kiadása, mely azt demonstrálja, hogy milyen jól különíti el a szoftverünk az azonos címmel rendelkező, de valójában eltérő műveket. A „Szerelem”, „Franciaország története”, „Szeged” kifejezések olyan általános címek, melyeket az idők során már több mű is megkapott.

A <http://konyvtar.info/utca/katalogus.php> címen ezeket, s még néhány további művet sorolunk fel. Az adott címmel rendelkező rekordokat művek szerint csoportosítjuk, így sok-sok találat helyett csak annyi lesz a katalógusunkban, ahány művet beazonosítunk. A művek természetesen kiadások szerint elkülönítve is megjeleníthetők, valamint mű és kiadás szintjén is lekérdezhető a könyvtári státusz az összes MOKKA, ill. ODR-lelőhelyen egyszerre. Az ízelítő csak néhány címet mutat, de ha küldenek nekünk továbbiakat, szívesen kitesszük azok feldolgozását is. A háttérben már minden olyan könyvtár állományát feldolgoztuk, amelyik a kutatási projektünket adataival támogatja. Adatokat szolgáltatott: a Berzsenyi Dániel Főiskola (Savaria Egyetemi Központ), a BME-OMIKK, Corvinus Egyetem, az ELTE, a MEK, a Magyar Tudományos Akadémia Könyvtára, a Miskolci Egyetem, az Országos Idegennyelvű Könyvtár és a Pécsi Tudományegyetem Könyvtára.

(Kardos András tájékoztatójából, Katalist, 2008. szeptember 9.)

web: <http://konyvtar.info/utca/> („Katalógus” menüpont)

prezentáció: <http://www.slideshare.net/andrisi/utca-szombathely/>

blog: <http://utca.klog.hu>