

FINNUGOR KULTURÁLIS ÖRÖKSÉG AZ INFORMÁCIÓS TÁRSADALOMBAN

Darányi Sándor – Dominich Sándor

a néprajztudomány kandidátusa, osztályvezető,
Oktatási Minisztérium K+F Helyettes Államtit-
kárság, EU Tudományos és Technológiai Együtt-
működési Osztály – sandor.daranyi@om.hu

a matematikai tudomány kandidátusa,
egyetemi docens, Veszprémi Egyetem,
Számítástudomány Alkalmazása Tanszék –
dominich@dcs.vein.hu

Bevezető

2001. szeptember 14-én rendezte az Országos Széchényi Könyvtár és a Finnugor Világkongresszus Magyar Szervezete a *Kulturális innováció és információkeresés: A finnugor kulturális örökség digitalizálása és automatizált tartalmi feltárása* című konferenciát (Darányi és Dominich, 2001), finnugrisztika és informatika közös témájában. Az alábbiakban az előadások motivációival, tudományos hátterével és perspektíváival foglalkozunk.

A gyűjtéstől a digitalizálásig

A kis finnugor népek sorsát szívükön viselő kutatók nemrégiben még az anyaggyűjtésre fektették a hangsúlyt. Ma is sürgető feladat az értékek konzerválása, illetve az adattárakban felhalmozódott gyűjtemények átfogó, olcsó publikálása (UNESCO, 1989, 1999; Európai Tanács, 1993).

Minden nép fennmaradása szempontjából igen fontos, hogy az anyanyelv ne szoruljon be a szűk családi keretek közé, ne maradjon meg a szóbeliségben, hanem minél több területen, például az iskolai oktatásban, a közigazgatásban is jelen legyen. A finnugor kulturális örökség digitalizálása nemcsak azért fontos, hogy számos veszélyeztetett kultúra fennmaradjon az utókor számára, hanem azért is, mert e nélkül a finnugrisztika tudományterülete menthetetlenül lemarad

más tudományok mögött, amelyek már most használják az informatika lehetőségeit.

A digitalizálástól az adatbázisokig

A digitalizálás a kulturális örökséget hordozó szimbólumgyűjtemények létrehozása, fenntartása, kezelése és tartalmi feltárása. Az internetre tett adatbázisok értéke azért nagy, mert a legközvetlenebb módon a terepmunkából származnak, tehát a gyűjtő és a közreadó kapcsolatai a legszorosabbak, ugyanakkor multimodálisak is: a világháló természetéből adódik, hogy szöveges, képi és hangzó információtjóról egyszerre olvassunk, látunk, hallunk, mintegy jelen lehetünk rokonaink életében.

Finnországban már tizenöt nyelvészeti adatbázisa van a kis finnugor nyelveknek, készül a *Lapp Enciklopédia*, illetve az ELTE Finnugor Tanszékének rokon ihletésű terve, a *Finnugor Enciklopédia*. Ki kell emelni az MTA Nyelvtudományi Intézete által koordinált *Uráli Morfológiai Elemző* projektet, illetve a Finn Irodalmi Társaságnak a finn ösköltészet digitalizálására indított projektjét.

Kultúrtér és kvantifikálás

E nagy munkák jóvoltából új lehetősége nyílik annak, hogy a kultúrtérhez természet- és műszaki tudományos látószögből (is) közelítsünk és azt ilyen eszközökkel (is) képesek legyünk vizsgálni. Kiderül ugyanis, hogy a

kvalitatív jellegűnek hitt jelentéstér kvantitatív módon is vizsgálható. A kvalitatív tér kvantitatív analízise hasonlít Gödel logikavizsgálatához számokkal (Kneale és Kneale, 1962), ám a kultúrtérben jóval több kategóriával kell szembenézni.

A közgyűjteményekben megtestesülő szimbólumkincs a számítógépes adatbázisokban áll először össze azzá az egésszé, amelyben az emberi jelhasználat természetrajza a maga tisztaságában, laboratóriumi körülmények között reprodukálható. A számítógéppel olvashatóvá tett kulturális örökség adatbázisai azok a laboratóriumok, amelyekben kísérletszerűen, tehát az eredmények ismételhető, előrejelző (prediktív) és utánozható (modellezhető) vonásaira összpontosítva vizsgálhatjuk egy fontos találkozási pont, bölcsészet és természettudomány határterülete viselkedését.

Az eredmények implikációi

Néhány előadó nem adatállományok közzétételével, hanem azok elemzésével foglalkozott. Kísérleteikben a nyelvi jelentés hordozója nem a fonémikus vagy grafémikus jel, tehát a hang vagy a betű, hanem a geometria, a látott tér. Mivel módszerük, az automatikus osztályozás nem három-, hanem sokkal több dimenziós tereket hoz létre, ezek vizualizálása a maguk teljességében észlelésünk korlátaiba ütközik. A korlát azonban megkerülhető azzal, hogy e tereket mégis három kiterjedés mentén, domborzat gyanánt, azaz mezőként ábrázoljuk úgy, hogy a mező hosszúsága, szélessége és magassága a szemantika bizonyos (mérhető) komponenseinek felel meg. E megközelítés – amely *Jost Trier* szemantikai mezőfogalmát állítja párhuzamba a fizikaival – óhatatlanul veti fel: létezhet-e a nyelvi jelentésnek valamiféle formális elmélete, amely alapján következetesen kettéválaszthatnánk a tartalmat és a formát, a hordozottat meg a hordozót (Machlup és Mansfield, 1983). Az auto-

matikus osztályozás eredménye híven reprodukálja a nyelvi közmegegyezést, a jelentésviszonyokat (Ogden és Richards, 1930; Sebeok, 1994). Ugyanakkor a geometriai jelentéshordozó modelljét eddig Ludwig Wittgenstein kontextuális, Rudolf Carnap logikai szemantikájának, John Lyons szintetikus magyarázatának, valamint Trier elképzelésének sikerült megfeleltetni (Darányi, 2000). Ebből a szempontból jelentős előrelépés, hogy a közelmúltban a Johann Gottlob Frege, illetve John Saunders Peirce nevével fémjelzett ún. szemiotikai háromszög is értelmezhetővé vált a vektortér-modellben. A szemiotikai háromszögben a nyelvi jel, tárgya (jelölete) és fogalmi jelentése megbont-hatatlan hármasságot alkot úgy, hogy az emberi elmében tárolt jelentés egyben egy következő jel is, tehát talppontja a következő háromszögnek, s a jelezés (szemiózis, jelfolyam) e hármasságok láncolatában szövi át az élővilágot. Az új eredmény lényege az, hogy valós számok vektortere helyett komplex számokat használ, ezzel lehetővé válik az utalás, amelyben valaminek az a lényege, hogy nem önmaga, hanem másvalami. Ezt úgy érjük el, hogy az információt, a formai elemet a komplex szám valós koordinátájának feleltetjük meg, a tartalmat pedig a képzetes koordinátának. A teljes komplex koordináta ekkor értelemszerűen a tartalom és a forma egységét, közös helyét rögzíti a vektortérben.

Interakciós információkeresés

A jelentésviszonyokat rögzítő adatbázisokban az információkeresésnek jelenleg három klasszikus (Boole-féle, vektortér, valószínűségi), néhány alternatív (cluster, fuzzy, LSI, neurális hálózat, genetikai algoritmus, tudásbázis alapú), valamint két nem klasszikus (logikai, interakciós) modellje létezik.

Az interakciós modell (Dominich, 1994, 2001) a kvantummechanikai mérés koppenhágai értelmezésén (Bohr, 1928; Roland,

1994) alapszik. Az adatbázisban tárolt anyag adja a megfigyelt rendszert, a keresőkérdés pedig a mérőeszköznek felel meg. A keresőkérdés megválaszolását jelentő információkeresés a mérés folyamatának analógja: amiként a kvantummechanikában a mérés eredménye a mérőeszköz és a mérendő rendszer – nullára soha nem csökkenthető – kölcsönhatásának eredménye, a keresőkérdésre kapott válasz a kérdés és az adatbázis interakciójából bontakozik ki. A kérdés kölcsönhatásba lép az adatbázisban tárolt anyaggal, abba beépül, e beépülés hatására az adatbázis részlegesen átszerveződik, s az aktivált lokális memóriák adják meg a válaszokat a kérdésre. Az interakciós információkeresés szabványos tesztkollekciókon elért kísérleti eredményei, és implementációinak, gyakorlati alkalmazásainak (World Wide Web meta-kereső) keresési hatékonysága jóval meghaladja minden más modell teljesítményét.

Az interakciós modellnek elméleti jelentősége is van: minden klasszikus modell (és így minden alternatív modell is) e modellnek valamely sajátos esete egy paraméter függvényében (miközben maga az interakciós

modell önálló, az összes klasszikus és másik nem klasszikus modelltől különbözően és azoktól teljesen függetlenül is létezhet). Ez az eredmény az információkeresés modelljei fejlődésének belső konzisztenciájára és dinamizmusára utal, amit formális matematikai bizonyítás is igazol.

Összegzés

A finnugrisztika és a digitalizálás, illetve informatika összefogása mindössze szemlélteti azt, mire mehet egymással bölcsészet és természettudomány. Az egymástól hagyományosan távolinak minősített két terület, a kulturális örökség valamint a természet- és a formális tudományok kölcsönhatása egyben az egymásra utaltság egyik neme is, s ha érdemi lépéseket kívánnak tenni, akkor ki kell mozdulniuk sok éve birtokolt „felségvizeikről”, közösen felderítendő tengereken is kell hajózniuk.

Kulcsszavak: *kulturális örökség, digitalizálás, információ-visszakeresés, automatikus osztályozás, információs társadalom, finnugor, kisebbségek*

IRODALOM

- Bohr, N. (1928). A kvantumposztulátum és az atomfizika újabb fejlődése. In: Jánosy, L. (1971). (szerk.) *Kvantummechanika*, 231–249, Akadémiai Kiadó, Budapest
- Darányi, S. (2000). A field approach to word semantics. *Semiotica* 128-3/4, 259–280.
- Darányi S. és Dominich S. (szerk.) (2001). Kulturális innováció és információkeresés: A finnugor örökség digitalizálása és automatizált tartalmi feltárása. URL: <http://www.oszk.hu/ujdonsag/finn/index.html>
- Dominich, S. (1994). Interaction Information Retrieval. *Journal of Documentation*, **50(3)**, 197-212.
- Dominich, S. (2001). *Mathematical Foundations of Information Retrieval*. Kluwer Academic Publishers, Dordrecht, Boston, London
- Európai Tanács (1993). A bécsi egyezmény szövege megtalálható az alábbi címen: URL: http://www.age-of-the-sage.org/vienna_declaration.html#Vienna_Declaration_text
- Kneale, W. és Kneale, M. (1962). *The Development of Logic*. Oxford University Press
- Machlup, F. és Mansfield, U. (Eds.) (1983). *The study of information: interdisciplinary messages*. New York: Wiley
- Ogden, C. K. és Richards, I. A. (1930). *The meaning of meaning: a study of the influence of language upon thought and of the science of symbolism*. London: Kegan Paul, Trench, Trubner & Co.
- Roland, O. (1994). *The interpretation of quantum mechanics*. Princeton University Press, Princeton, New Jersey
- Sebeok, Th. A., ed. (1994). *Encyclopedic Dictionary of Semiotics I-III*. Berlin: Mouton de Gruyter
- UNESCO (1989). Ajánlások a hagyományos kultúra és a folklór védelmére (UNESCO Közgyűlés, 1989). EFI Communicationes 1. Bp., Európai Folklór Intézet
- UNESCO (1999). Cselekvési terv a szellemi kulturális örökség megóvására és újjáélesztésére (UNESCO Konferencia, 1999). EFI Communicationes 9. 1999. Budapest, Európai Folklór Intézet