

A humán tudományok és a gépi intelligencia.

Szerkesztette: Tolcsvai Nagy Gábor.

Gondolat Kiadó, Budapest, 2018. (200 lap)¹

1. A humán tudományokról, a bölcsész szakmákról még ma is sokan azt tartják, hogy a kutatóknak a saját fején kívül nincs szüksége egyébre, mint legfeljebb papírra és ceruzára. Ezt a felfogást olykor még a tudománnyal foglalkozók körében is tapasztaljuk. Ennek a tévhitnek az eloszlataához is hozzájárulhatott az a konferencia, amelyet az MTA I. Osztálya 2017-ben a Magyar Tudomány Ünnepeén rendezett meg, s amelynek anyaga Tolcsvai Nagy Gábor szerkesztésében, a Gondolat Kiadó gondozásában „*A humán tudományok és a gépi intelligencia*” címmel jelent meg, a Kertész András osztályelnök szerkesztette „*A Humán Tudományok Alapkérdései*” sorozat harmadik köteteként.

E sorozat célja, hogy az I. Osztály által képviselt tudományágak tudományelméleti kérdéseit, valamint más tudományokhoz fűződő viszonyát, illetve e tudományok társadalmi szerepét vizsgáló, magas tudományos színvonalú tanulmányokat tegyen közzé közérthető, a szélesebb közönség számára is hozzáférhető módon.

2. A kötet a szerkesztő, Tolcsvai Nagy Gábor tanulmányértékű előszavával (9–16) együtt tíz írást tartalmaz: ötöt az alapkutatások köréből, egy-egy tudományterület idevágó problematikáját elemezve. A nyelvtudomány helyzetét Prószéky Gábor mutatja be (17–32), az irodalomtudományét Debreczeni Attila (48–61), a klasszika-filológiáét Kozák Dániel (62–91), a néprajztudományét Nagy Károly Zsolt (92–126), a digitális bölcsészetét pedig Palkó Gábor (33–47) tekinti át. Az alkalmazott tudományok szemléletmódja érvényesül Németh Géza írásában a gépi beszédkeltésről (127–155) és Laki László János munkájában a gépi fordításról (156–183). Arról, hogy a számítógép hogyan segítheti az ember művészi alkotó tevékenységét, két művészeti ágban tájékozódhatunk: Szűts Zoltán tanulmányából a szépirodalmat (184–194), Vidovszky Lászlóéból pedig a zenét (195–200) érintően.

Meg kell említeni, hogy a számítógépes kultúra nemcsak az itt felsorolt tudományokra, hanem magukra az e tudományok által vizsgált jelenségekre is meghatározó erejű hatással van. Elég csak a nyelvnek mint kommunikációs eszköznek a digitális világ által kiváltott mélyreható változására gondolni, amely persze együtt jár az emberi közösség működésének a módosulásával is. Am e kötet ezeket a hatásokat legfeljebb csak áttételesen érinti, mivel figyelmét arra a kérdésre összpontosítja, hogy a számítógép alkalmazása miben nyit új távlatokat, és milyen problémákat vet fel a tudományos kutatásban a humán tudományok terén.

3. A technikai eszközök használata a humán tudományok körében a legrégebbi múltra talán a néprajztudományban megy vissza. A kép- és a hangrögzítés fontossága, azaz a kutatás multimediális környezete komoly motivációt adhat a digitális kultúra elterjedéséhez. Nagy Károly Zsolt azonban úgy látja, hogy – noha a kutatók gyakran használják a digitális technológiát – ez a folyamat a néprajztudományban mégis reflektálatlanul zajlott le. Szerinte

¹ Az MTA I. Osztálya által 2018. november 26-án rendezett könyvbemutatón elhangzott előadás szerkesztett változata.

e szakmában a számítógép olykor nem egyéb, mint „a modern enteriőrök reprezentatív kelléke, méregdrága levélnehezék” (121).

Ez a más tanulmányokban is felbukkanó szkeptikus, önironikus szemlélet – amely általában sem idegen a tudománytól – azonban leginkább a még erőteljesebb megújulás iránti igényt fejezi ki, s nem jelenti azt, hogy az adott tudományterületeken ne születtek volna már eddig is számottevő eredmények a digitális kultúra támogatásával. Nagy Károly Zsolt is említ ilyeneket a néprajztudomány köréből, én ezek közül itt most csak egyet emelek ki.

2011-ben megjelent munkájában Borsos Balázs a *Magyar néprajzi atlasz* anyagát dolgozta fel számítógépes klaszteranalízis segítségével. Az egyes jelenségeket matematikai-statisztikai módszerekkel vetítette egymásra a térinformatika eszközeit is felhasználva, és így a kultúra szerveződésének olyan térbeli mintázatát tudta felvázolni, amelyet hagyományos módszerekkel a tudomány korábban nem volt képes megoldani.

4. A számítógép szerepének az itt említett ellentmondásos megítélése a humán tudományokban azzal is összefüggésben van, hogy a számítógép sokféle módon van jelen a tudományban a szövegszerkesztő általánosan elterjedt funkciójától kezdve az „intelligens” gépként való felhasználásáig. Hogy milyen alkalmazásában tekinthető ez az eszköz a metaforikus antropomorfizálás révén ráruházott „intelligens” tárgynak, abban távolról sincs egyetértés a tudományban, és a kötet szerzői is sokféleképpen értelmezik ezt a problémát.

A kérdéskört legalaposabban Prószéky Gábor járja körül, aki szerint a gépi intelligencia éppúgy definiálhatatlan fogalom, mint ahogyan az emberi intelligencia is az. Ez utóbbi igen összetett jelenség, lényege a bonyolult szervezettség és az ebből adódó problémamegoldó képesség. A legnagyobb óriás számítógépek is legfeljebb 1%-át tudják szimulálni az emberi agy teljesítményének. Noha a tárhely mérete már az évtizedekkel ezelőtti C-64 gépek esetében is összehasonlíthatatlanul nagyobb volt az emberi memóriáénál, a gépi intelligencia fejlődésében annak van igazán nagy szerepe, hogy a mai számítógépek autonóm tanulási képességgel rendelkeznek, és így alkalmazkodni tudnak a környezetükhöz.

5. A számítógép hatalmas memórialehetőségeit az elsők között egy nagy hagyományú tudományág, a klasszika-filológia használta fel. Kozák Dénestől megtudjuk, hogy már az 1940-es években létrehozták az IBM segítségével Aquinói Szent Tamás terjedelmes életművének konkordanciáját, az 1970-es évekre pedig elkészült egy 27 millió szövegszót tartalmazó görög digitális korpusz is, amely ma már több mint 100 millió szövegszót tartalmaz. A klasszika-filológia lényegében zárt korpuszal dolgozik, ám hagyományos módszerekkel a legszélesebbkörűen tájékozott tudós sem képes átlátni ennek az anyagnak az elenyésző töredékét sem. A számítógép alkalmazása főleg az intertextualitás vizsgálata terén hozott fontos eredményeket. Az elme által átfoghatatlan szövegek együttes „olvasása” új stratégiákat teremtett meg: a „távoli olvasás”, az „algoritmikus olvasás” révén a kvantitatív elemzésre került a hangsúly, s ez új összefüggések felismerését tette lehetővé.

6. A digitális szövegtörzsek mennyiségének hatalmas méretű gyarapodása a gépi fordítás területén is átütő változást hozott. Prószéky Gábor szerint a nyelvtudomány első találkozása a számítógéppel éppen ezen a területen történt meg. A számítógép fordításra való felhasználása tulajdonképpen egykorú a számítógép létevel: a második világháború idején katonai titkosírások megfejtésére alkalmazták a gépeket, és az eredmények nagy reményeket keltettek a gépi fordítás megvalósíthatóságát illetően. A korai szabályalapú fordítórendszerek a szövegből kinyerhető lexikális, grammatikai, szemantikai szabályok alapján működtek, ám meglehetősen gyengén, noha időközben a számítógépek teljesítménye tetemesen növekedett.

Új perspektívát teremtett az adatvezérelt, statisztikai alapon fordító rendszerek kidolgozása. Ennek háttérében az áll, hogy egyre nagyobb a párhuzamosan lefordított és digitálisan tárolt szövegek korpusza. A fordítás hatékonyságát nagyban befolyásolja, hogy mely nyelvpárok között történik: a magyar nyelv számára az indoeurópai nyelvekétől eltérő tipológiai sajátosságai nem teremtenek kedvező helyzetet a gépi fordításban. Laki László János nem minden alap nélkül szkeptikus a jövőt illetően: „nehéz megmondani, hogy [...] bármikor is sikerül-e létrehozni tökéletes fordítórendszert” (180).

7. Prószéky Gábor szerint a strukturalista nyelvészet – mivel kialakulásakor jórészt a formális logikán alapult – szinte predesztinálva volt a számítógépekkel való találkozásra, ami azonban valójában nem történt meg. A korpusznyelvészetet viszont maga a digitális kultúra hozta létre, s mint tudományos irányzat – a fentiekből következően némileg paradox módon – szinte ellenpólusa a strukturalista nyelvészetnek.

A számítógépes nyelvészeti elemzések csillapíthatatlan adatéhsége felmérhetetlen mennyiségű szöveget von be a kutatásba. Ennek hasznát Prószéky a disztribúciós szemantika robbanásszerű fejlődésének példáján mutatja be. A vizsgálatban használt szóbeágyazási modell tulajdonképpen szövegkörnyezeti típusok hasonlóságát mutatja ki, és a gyakorisági elemzés révén igazolja azt az elméleti definíciót, amely szerint a jelentés a nyelvi jel használati szabályaként értelmezendő.

Prószéky Gábor szerint napjainkban egy újfajta gépi intelligencia van megszületőben, amelynek fontos jellemzője, hogy az emberi elme a kiinduló és a köztes adatok között nem lát összefüggést, így a gépi intelligencia egyre inkább önjáróvá válik, ezáltal pedig ebben a kölcsönhatásban megnő az emberi intuíció szerepe, főképpen a kiinduló paraméterek beállítása [terén](#). Prószéky úgy véli, hogy a gépi intelligencia mechanikus jellegének és az emberi intuíciónak az ötvözése vezethet leginkább sikerre.

8. A számítógép tevékenységének bizonyos mértékű kontrollálhatatlansága, a „számítógép mint fekete doboz” gondolata megjelenik Palkó Gábor írásában is. Ezt a metaforát leginkább az magyarázza, hogy a gép a használat során folyton újra konstruálja magát. Megítélése szerint legalább ennyire fontos körülmény az is, hogy a digitális kultúra átrendezi a kutatás tereit, mivel létrehozza a virtuális kutatókörnyezetet, s a kutatás tárgya, dokumentumai is másképpen formálódnak meg ebben a térben, mint annak előtte.

Palkó a kutatómunkának ezeket az új jellegzetességeit Esterházy Péter prózai szövegeinek hasonlósági, „stilometriai” vizsgálatával szemlélteti. A kutatási tér kitágulását jelzi már magának e nagy szövegkorpusznak az elemzése is, amelyre az egyéni rálátás egészében aligha valósítható meg. A digitalizálás magával hozza különböző szakmák együttműködését is: nyelvész, irodalmár, statisztikus, informatikus dolgozott együtt ebben a projektben. Palkó hangsúlyozza, hogy mindez egyúttal a kutatások interdiszciplináris jellegét eredményezi, és a szerzőfunkció megváltozását is magával hozza.

9. A beszédprodukciónak gépi létrehozása ugyancsak több szakma együttműködését igényli. Ennek az alkalmazott tudományterületnek az elméleti alapjait több mint 200 évvel ezelőtt Kempelen Farkas fektette le, a számítógépet pedig az 1950-es évektől vonták be a gépi beszédalkotás kutatási-műszaki feladatainak megoldásába.

Németh Géza bemutatja, hogy amíg csupán az artikulációs csatorna szolgált a modellezés alapjául, addig csak jellegzetes robotszerű beszédet sikerült produkálni. A hangsúly a 80-as évektől került át az akusztikai oldal feldolgozására, és a két megközelítés kombinálása egyre jobb eredményeket hozott, amit a korpuszalapú modellek tovább tökéletesítenek.

Ha belegondolunk, nincs túl távol az az idő, amikor magyar beszédünket egy beszédfelismerő program tökéletesen lejegyzí, ezt egy fordító program lefordítja pl. finnre, a finn írott szöveget pedig egy felolvasó program közvetíti finn beszédpartnerünknek, és viszont. Ma még nehéz azt megjósolni, hogy mindez vajon milyen hatással lesz például az idegennyelv-tanulási hajlandóságra, vagy szélesebb értelemben a globális kommunikációra.

10. A számítógép kreativitását az irodalom és a zene is igyekszik felhasználni: az első számítógéppel írt versek és zeneművek nagyjából egy időben, az 1950-es évek végén jelentek meg. Vidovszky László a zenét ebben a relációban meglehetősen redukált fogalomként értelmezi, hangok kombinációinak megvalósulásaként fogja föl. Csak sajnálni tudjuk, hogy a könyvformátum mint információhordozó nem adott lehetőséget számítógép alkotta zeneművek közzétételére.

Sajnos gépverseket sem olvashatunk Szűts Zoltán e témát tárgyaló írásában. A szerzőség problémája itt is jócskán felmerül: Szűts szerint a mű poétikai, esztétikai tekintetben leginkább a programszoftver szabályainak van alárendelve. Abból, hogy a számítógépes irodalom jóval kevésbé kap elismerést, mint például a gépi vizuális műalkotások, a szerző arra következtet, hogy az irodalom a legkonzervatívabb művészeti ág.

11. Debreczeni Attila, aki az irodalomtudomány és a digitális kultúra kapcsolatáról a kritikai kiadások kapcsán szól, úgy véli, hogy a jelenlegi digitális kiadások jórészt még csupán a nyomtatott kiadások imitációi, mégis, némi intézményesülésről is be tud számolni ezen a területen. Az ez ügyben érintett szakemberek ugyanis az Akadémia égisze alatt már a munkálatok kezdetén, csaknem másfél évtizeddel ezelőtt konszenzusra jutottak az együttműködést és ennek főbb alapelveit illetően. Debreczeni Attila szerint ez nem jelentheti azonban a teljes uniformizáltságra való törekvést, mert az alkalmazott keretrendszernek kellően rugalmasnak, az egyedi feladatokhoz igazodónak kell lennie.

Kiemeli, hogy a digitális kiadásnak az identitása is átértékelődik, mivel nem eredményként (mint a könyv), hanem folyamatként értelmeződik, és ennek megfelelően nem olvasói, hanem felhasználói vannak. E munkákban nemcsak a kritikai kiadások alapjául szolgáló szövegváltozatok egyenrangúsága valósul meg, hanem a szöveggel a kommentárok is egyenrangú, oda-vissza irányú kapcsolatban vannak. Ez pedig a kommentárok új műfaji típusát hozza létre, ami a kutatások számára korábban nem tapasztalt lehetőségeket teremt. A digitális kiadást Debreczeni Attila egyfajta laboratóriumi környezethez hasonlítja, ahol a textológus-filológus az informatikus szakemberekkel együttműködve új típusú kutatómunkát végez.

12. A kötet tanulmányai jól mutatják azt a helyzetet, amelyet a humán tudományoknak a számítógépes kultúrával való találkozása jelenleg mutat: szinte beláthatatlan, de ígéretes távlatok nyílnak meg a kutatás számára, ugyanakkor az első lépések megtétele sem könnyű ezen a terepen. A kötet szerzői éppúgy bemutatják ezt a kezdeti, néha kissé tanácstalan botladozást, mint ahogyan szólnak olyan nagyszerű eredményekről is, amelyek az új módszerek alkalmazása nélkül nem születhettek volna meg.

Debreczeni Attila úgy látja, hogy a kutatók ezen a téren még inkább a pidgin nyelv fázisában vannak, egyelőre meglehetősen távol attól, hogy ez a nyelv kreollá alakuljon. Lehet, hogy ez ma még valóban így van, előbb-utóbb azonban kétségkívül elengedhetetlen követelményé válik, hogy a humán tudományok művelői is anyanyelvükként használják a digitális-számítógépes kultúra nyelvét.

Hoffmann István
az MTA doktora