

A HUMÁN
TUDOMÁNYOK
ÉS A GÉPI
INTELLIGENCIA



A HUMÁN TUDOMÁNYOK ALAPKÉRDÉSEI

Az MTA Nyelv- és Irodalomtudományok Osztályának könyvsorozata

Sorozatszerkesztő

Kertész András

3.

A HUMÁN
TUDOMÁNYOK
ÉS A GÉPI
INTELLIGENCIA

Szerkesztette
Tolcsvai Nagy Gábor

Gondolat Kiadó
Budapest, 2018

A kötet megjelenését a Magyar Tudományos Akadémia támogatta.

Minden jog fenntartva. Bármilyen másolás, sokszorosítás,
illetve adatfeldolgozó rendszerben való tárolás
a kiadó előzetes írásbeli hozzájárulásához van kötve.





© Szerzők, szerkesztő, 2018

*www.gondolatkiado.hu
facebook.com/gondolat*

A kiadásért felel Bácskai István
Szöveggondozó Gál Mihály
A kötetet tervezte Lipót Éva
Tördelő Sörfőző Zsuzsa

ISBN 978 963 693 870 3
ISSN 2560-0885

Tartalom

Sorozatszerkesztői előszó	7
Tolcsvai Nagy Gábor	
Előszó	9
Prószéky Gábor	
A nyelvtudomány találkozásai a gépi intelligenciával	17
Palkó Gábor	
A digitális bölcsészet kultúrtechnikái. Virtuális kutatókörnyezetek	33
Debreczeni Attila 	
Az elektronikus kiadás színe és visszája	48
Kozák Dániel 	
Digitális szövegtörzsek a klasszika-filológiában	62
Nagy Károly Zsolt 	
Interaktív etnográfia. Néprajzi kutatás interfészek közti terepen	92
Németh Géza 	
Kempelentől a WaveNet-ig: a gépi beszédkezelés tudományának fejlődése	127
Laki László János	
Mesterséges intelligencia a gépi fordításban	156

6 ♦ TARTALOM

Szűts Zoltán

Irodalmi Nobel-díjas mesterséges intelligencia?
A gépversektől a Big Data rövid történetekig 184

Vidovszky László

DEUS EX MACHINA – a tehetséges számítógép 195

Sorozatszerkesztői előszó

Az MTA Nyelv- és Irodalomtudományok Osztálya 2014-ben a humán tudományok alapkérdéseivel foglalkozó konferenciasorozatot indított. A könyvsorozat megjelentetésének célja az osztályunkban képviselt tudományágak tudomány- és ismeretelméleti alapkérdéseit, valamint társadalmi funkciójának és a természettudományokhoz való viszonyának problémáit elemző, a konferenciákon elhangzott előadásokra épülő tanulmánykötetek közreadása.

Ily módon a sorozat kötetei e kérdésekre adnak szisztematikusan kifejtett, magas tudományos színvonalú, ugyanakkor közérthetően, a szűk szakmán túl szélesebb olvasóközönségnek is szóló válaszokat. Szándékunk szerint a kötetek szerzői – a mindenkor konferenciák előadói – szakterületük kiemelkedő művelői közé tartoznak, és tanulmányaik az adott kérdéskör fontos problémáit tárgyaló, autentikus írások lesznek.

A sorozat első kötetét abban a reményben bocsátjuk útjára, hogy a sorozat egy olyan szemléletmód fóruma lehet, amely a szerzőket és az olvasókat egyaránt a kutatómunkánk alapjaira való reflexióra, kristálytiszta gondolkodásra és a tudományos megismerés sokféleségében rejlő értékek felismerésére készíti.

Kertész András



Tolcsvai Nagy Gábor

Előszó

1. A konferencia célja

A bölcsészettudományok és az általuk vizsgált jelenségek körében a műszaki eszközök használata évezredek óta jellegzetesen ritka és nem jelentős. A számítógép megjelenése ezen a helyzeten nagymértékben változtatott. Az emberi elme és a gépi, de ember által működtetett műveletvégzés alkotó és kommunikációs kapcsolata, interakciója összetett és mindinkább kétirányú.

A gépi műveletvégzés hosszú évtizedek óta egyre nagyobb szerepet játszik a humán tudományok működésében, immár nem pusztán adattárolás vagy szövegszerkesztés és -továbbítás céljából. A nagy mennyiségű kutatási adatok objektív gyűjtése, különböző célú keresése, feldolgozása az ezredfordulóra megkerülhetetlen eljárások lettek. E tudománytörténeti folyamatban a kutatástervezés, a hipotézisek felállítása, a kutatási eredmények előrejelzése, ezáltal a tárgytudományos elméletek és módszertanok, illetve a tudományfilozófiai alapok is a gépi feldolgozás hatása alá kerültek és kerülnek, valamilyen mértékig. Mindezzel szoros kulturális összefüggésben a humán tudományok által kutatott jelenségek, a gyakorlat szintén erősen a gépi működés függvényévé vált, a mindennapi kommunikációtól kezdve egyes művészeti ágakig.

A Magyar Tudományos Akadémia I. Osztálya e kettősség viszonyában közelítette meg a humán tudományok és a gépi műveletvégzés kérdését 2017. évi konferenciáján, a Magyar Tudomány Ünnepeán.

A kettős tematika egyik fő témája a gépi műveletvégzés szerepe a tudományos tevékenységben. A fontos kérdések itt a kö-

vetkezők. Mennyiben eszköz és mennyiben a kutatás alakítója a számítógép műveleti alkalmazása. A műveletek végeztetése és az adatbázisok fajtái és működtetése hogyan alakítja a kutatás keretét, az adatok összetételét és rendszerét, mennyire önjáró a gépi intelligencia. Milyen fokú a kísérleti vagy modelláló műveletek beprogramozásának és végrehajtásának, illetve a korpusz létrehozásának és a korpuszban való irányított keresésnek az emberi és gépi irányítottsága. Miképp járul a számítógépes kutatás meg alapozásához a hagyományos kutatástervezés és -módszertan, illetve fordítva: hogyan alakítja át a gépi kutatás a humán feldolgozást; alakítja-e a tudományelméletet, és ha igen, hogyan. Hol mond csődöt a mesterséges intelligencia a kutatásban. A fejezetek az irodalomtudomány, a nyelvtudomány, néprajztudomány és a klasszika-filológia köréből tárgyalják a kérdéseket.

A konferencia tematikájának másik összetevője a gépi intelligencia szerepe a kreatív tevékenységben. Miben segíti a gépi tevékenység az emberi alkotótevékenységek ezredforduló utáni rendszerét. Ebben a részben van szó a számítógépes fordításról, az emberi jellegű beszéd gépi előállításáról, művészet és gépi intelligencia kapcsolatáról a számítógépes (irodalmi) írás és a számítógépes zeneszerzés területén.

2. A gép és a kultúra

A gép, a műszaki jelenlét az emberi kultúra meghatározó része. A gép nem természeti, és nem is humán, miközben a természeti felhasználásával ember által készül. A technika a kísérletező szubjektum alkotótevékenységével, a hatalomhoz és a szabadsághoz való viszonyával áll összefüggésben.¹ A technika, pontosabban a fizika innen nézve a megismerés mint a megismerés történő előállítása mutatkozik meg. A technika a *van* helyett a *kellre* irányul: a meglévő a még meg nem lévő megragadását teszi lehető-

¹ Vö. Blumenberg 2015: 47.

vé.² A folyamatban a *hogyan* kerül előtérbe, amely a munkában mutatkozik meg. Az ember ezért nem annyira műszaki termékek létrehozója, hanem inkább saját magát műszakilag, technikailag megvalósító lény.³ Ez az ontológia az ember világot alakító vagy alkotó, Isten helyébe lépni kívánó szerepét mutatja, nyilvánvalóan a maga kérdéseivel és korlátaival.

A gépnek – tágabban és általában, illetve az ezredfordulón a számítógépnek – számos általános értelmezése azonosítható. Ezek a magyarázatok a tevékenykedő ember történeti szociokulturális világában helyezik el a gépet.

A gép például szerszám. Ebben az esetben a gép az ember által létrehozott segítő tárgy, amely a humán erőnél nagyobb behatást képes kifejteni, emberi működtetéssel. Így szerszám az egyszerű mechanikus számológép, amely a mennyiségeket fizikailag megjeleníti.

A gép más esetben eszköz. Az eszköz nem tesz semmit, azzal tesz valamit a dolgozó ember. Galilei távcsöve eszköz, amellyel alkotója szabad szemmel nem vagy nem jól látható égitesteket tudott vizsgálni, és e vizsgálatokból új következtetéseket tudott levonni. A tárgyi eszköz ezáltal megváltoztatta az ember világbeli helyének értelmezhetőségét, mert visszavonta az ember középponti helyét az Isten által teremtett világegyetemben, a középkori kozmikus topográfiában (vagy pontosabban: megadta a lehetőséget a topográfiai centrális hely felváltására).⁴

A gép, főképp már a számítógép „magától” is tesz valamit, parancsra, és program alapján, mintegy belső működéssel külsőleg megnyilvánuló cselekvést hajt végre, előzmények (program és parancs) nyomán. E ponton jelenik meg az animálás, a metaforikus antropomorfizálás: a gép biológiai létezőként képződik le, azaz a számítógép mesterséges agy. A robotok gyakran kapnak emberi vagy háziállat alakú fizikai szerkezetet és külsőt, képen is megjelenítve a humán analógiát. Így a gép például szolgál,

² Vö. Blumenberg 2015: 49.

³ Vö. Blumenberg 2015: 49.

⁴ Vö. Blumenberg 2015: 60–63.

amely zokszó nélkül végzi feladatát. Vagy társ a napi tevékenységben (ez kivált érvényes a főképp okostelefonokkal működtetett közösségi terek elektronikus kommunikációjában). Vagy a fölérendelt abszolút (megbízható) cselekvő, a mindent megoldó, mégsem személyes lény, a kultusz tárgya és egyúttal alanya. Ekkor a számítógép „titokzatos” entitás, nem szubjektum, nem ittlét (Dasein), nincs öntudata, bár a róla való beszédben gyakran így konceptualizálódik. Itt jelenik meg a géppel kapcsolatos teleológia, a jövőre irányuló célelvű megjelenítés: átveszi az ember munkáját, sőt azt meghaladja képességeiben, amikor megvalósítja a jövőt.

Az emberi intelligencia viszonyítási pontja korábban egyöntetűen Isten, de a létező természeti világ középpontjában mégis az ember állt. Ezzel szemben a felvilágosodás az ész mindenhatóságát jelentette ki; e felfogás szerint az ember került a világ Istent tagadó középpontjába mint az e világot megismerő, „a természetet leigázó”, a világot átalakító legfelső lény. A gép ezt a racionalista világméretet megzavarja, hiszen az ember létrehozott egy olyan szerkezetet, amely működésében jelentős paraméterekben meghaladja az emberi teljesítményt (ilyen a memória kapacitása és pontossága, az adatok hozzáférhetőségének állandósága, a működés gépies ismétlődő azonossága), más paraméterekben viszont nyilvánvalóan alulmúlja, illetve e paraméterek idegenek a géptől (ilyen a kreativitás, a képzelőerő, az újítás, az elkötelezettség és az etika). Mindennapi példákat említve: a korlátlanak tűnő lehetőségek kihasználásán alapul a videojátékok előállítás és a velük való foglalatosság, a gépi sakk vagy az önvezető kocsik, vagy akár a beszédgenerálás és az emberi beszéd megértése

3. A mesterséges intelligencia

A mesterséges intelligencia a természetes intelligenciával szemben vagy vele viszonyban van meghatározva, nemcsak a köznapi beszéd metaforáiban, hanem gyakran a kutatásban is. A mesterséges intelligencia gép, amely eszköz külső ingert észlel, feldol-

goz és válaszol rá.⁵ Itt egyértelműen a professzionális (nem személyi) számítógépről, gépek összekapcsolt működéséről van szó. A mesterséges intelligencia emberi létrehozásának több forrása azonosítható: az egyik az önmagát meghaladni kívánó transzcendencia utáni humán vágy, a másik a kíváncsiság és a képzelőerő, a harmadik a megismerő és munkatevékenységek észszerű szervezése és a teljesítménynek az emberi képességeket meghaladó fokozása.

Azonban éppen az intelligencia fogalmának játékba hozása miatt kérdés, vajon a gép tudja-e teljesíteni például a következő emberi műveleteket: észlelés, tanulás, érvelés, tervezés, mozgás és tárgymanipluláció. Még inkább kérdéses a kreativitás vagy társas kapcsolatokra képesség megléte. A számítógép programozása mesterséges, logikai alapú nyelveken történik. Ez nem felel meg a természetes emberi nyelvnek, és csak rokonítható humán elmemodellekkel, vagy éppen a különbségek a szembeűnők. A számítógép nem része valamely autopoietikus kommunikációs rendszernek.⁶ Föltehetőleg nincsen emberi megismerés közösség, kultúra és kommunikáció nélkül, a gépek pedig erre nem képesek, ezért a humán párhuzam itt is csorbát szenved.

Az ember testben létezik, a test fizikai, biológiai és társas fennmaradásáról gondoskodnia kell. Ehhez külső és társas forrásokra van szüksége. Ebben az értelemben függősége fokozat szerinti, az öngondoskodásnak van lehetősége. A gép függő, függ a hardvertől (ez rajta kívül áll), az energiabeviteltől (ezt magának nem tudja biztosítani), a programozástól (csak kezdeti lépések után lehetséges korlátozott önállósága).

Mivel a számítógép ember alkotta műszaki szerkezet, folyamatosan jelen lévő tényező a műszaki lehetőségek hatóköre, a kezelhető adatok mennyisége és a számítógép tényleges, a mindenkori lehetőségek határáig elvitt teljesítménye.

⁵ Vö. például Poole–Mackworth 2017; Russell–Norvig 2009.

⁶ Ehhez lásd Luhmann 1998.

A mesterséges intelligencia ember által tervezett folyamat eredménye.⁷ Lehetséges, hogy olyan önalkotó alakulástörténetté válik, amelynek egyes pillanatnyi eredményei csak részben láthatóak előre, és részben megjósolhatatlan új kérdéseket és fejleményeket nyitnak meg.

4. A számítógép és a tudomány

A tudomány a számítógéppel sajátos viszonyban áll. Egyrészt a tudomány hozta létre a gépet, az elméleti modellezésektől a fizikai megvalósítás kidolgozásáig. Másrészt a tudomány a számítógéptől több szerep ellátását várja: az adatok maradéktalan, hiba nélküli tárolását, kezelését, program szerinti feldolgozását, illetve az adatok alapján eredmények előállítását, folyamatok modellálását, vagyis nagy adattömegekkel olyan, program szerinti műveletek végrehajtását és olyan „kreatív” tevékenységet, amelyek meghaladják az egyedi vagy csoportos emberi kutatói kapacitást, képzelőerőt, és betöltik a heurisztika vagy az új nézőpont, elméleti kezdeményezés esetleges emberi hiányát.

Erős bizakodás és visszafogott borúlátás egyaránt megmutatkozott, és ez ma is a helyzet. A bonyolult számítógépes rendszerek működésében kutatók gyakran látják a megoldás forrását például egyes betegségek gyógyításának ügyében (például a diagnosztika vagy a mikrosebészet területén okkal), miközben a szkeptikusok továbbra is az ember szellemi tevékenységében bíznak, bármilyen korlátokkal küzd is az. Az önvezető kocsik jelen idejű példája jól mutatja e kettősséget: az egyik oldalról megképződik a boldog jövő, amikor a gyors és biztonságos közlekedést gépek működtetik, miközben a másik oldalról súlyos problémák jelentkeznek a sikeresnek tűnő kezdeti szakaszok után.

A számítógép tudományos alkalmazása azt mutatja a felhasználónak, hogy a gép struktúra, szerkezet, amely működik.

⁷ Lásd Nillson 2009; Pléh szerk. 1998.

Egyúttal rendszer, amely műveleteket hajt végre külső hatásra (program és parancsra), a többség számára nem minden részletében átlátható módon. Vajon „intelligens”-e ez a gépi működés a legösszetettebb és legnagyobb adatmennyiséget megmozgató folyamatokban? A jelen kötetben olvasható tanulmányok óvatos és tartózkodó válaszokat adnak e kérdésre. E válaszok rokonságot mutatnak John Searle nevezetes írásának⁸ álláspontjával: az emberi gondolkodás intencionális jellegének szimulációja nem tekinthető megértésnek. A gépi programozás és műveleti tevékenység bizonyos pontokon „önállósul” abban az értelemben, hogy az azt irányító kutató már nem tudja minden részletében követni a gépi fejleményeket, de ez nem a gép kreativitásának az eredménye (ezt példázza kötetünk első tanulmánya a nyelvtudományról szólva). Vagy a gépi fordítás és a gépi beszéd-előállítás itteni áttekintéseiben azzal szembesülhet az olvasó, hogy különböző elméleti és módszertani alapon álló programokkal kísérlik meg az emberi mentális tevékenység egy-egy területét mesterségesen előállítani, javuló, de nem tökéletes megoldásokat eredményezve.

Ugyanakkor az itt olvasható tanulmányok is arról tanúskodnak, hogy a számítógép a kutatási lehetőségek korábban nem sejtett gazdagságát nyitja meg. Olyan nézőpontokat, összefüggéseket tesz hozzáférhetővé, amelyek a tudomány megelőző történeti szakaszaiban meg sem formálódtak, új elméleti és módszertani belátásokat eredményezve. Mindez az adott humán tudományok tárgyainak megértési feltételeit jelentősen átalakítja, főképp szélesíti. Az látható tehát, hogy új tematikai és feldolgozási lehetőségek és egyúttal korlátok jelentkeznek a humán tudományokban. A jelen kötet kiváló bemutatása a gépi adatfeldolgozás lehetőségeinek és mélységének, a rendkívül gazdag eredményekkel és beláthatóság visszakérdezésével. Emellett az emberi alkotótevékenység gépi generálással kiterjesztése olyan összefüggések (például zenei vagy nyelvi viszonyok) hozzáférhetővé tételét valósítja

⁸ Vö. Searle 1996.

meg, amelyek eddig földérintetlenek maradtak, bár algoritmikus alapon, föltehetőleg nem az emberi gondolkodás analógiájára. S jóllehet a tudomány egyetemes, a bölcsészettudományok tárgya eltér a természettudományokétól, ezért az emberi természet megértésének számítógépes leírása, modellálása például a természeti törvények által működő fizikai világ kutatásától részben eltérő gépi eljárásokat kíván meg.

A számítógép és a tudomány kapcsolata nyitott világ. Hogy miképp fog vele élni a tudomány a későbbiekben, miképp fogja mindez alakítani az elméletképzést és annak nyomán az adatok adattá nyilvánítását és értelmezését, a rendszerépítést, és hárátható lesz-e a módszertan talán túlzó előtérbe kerülése – ez a jövőbeli folyamatok egyik fő kérdése. Ezeket a most alakuló irányokat mutatják be a jelen kötet tanulmányai a maguk területén.

Irodalom

- Blumenberg, Hans 2015. Technik und Wahrheit. In *Schriften zur Technik*. Frankfurt am Main: Suhrkamp.
- Luhmann, Niklas 1998. *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Nilsson, Nils 2009. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. New York: Cambridge University Press.
- Pléh Csaba (szerk.) 1998. *Megismeréstudomány és mesterséges intelligencia*. Budapest: Akadémiai Kiadó.
- Poole, David – Mackworth, Alan 2017. *Artificial Intelligence: Foundations of Computational Agents*. Cambridge: Cambridge University Press.
- Russell, Stuart J. – Norvig, Peter 2009. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall.
- Searle, John R. 1996. Az elme, az agy és a programok világa. In Pléh Csaba (szerk.): *Kognitív tudomány*. Budapest: Osiris – Láthatatlan Kollégium. 136–151.

Prószéky Gábor

A nyelvtudomány találkozásai a gépi intelligenciával

1. Bevezetés

Ha az intelligenciát nem is könnyű definiálni, az agy tárolási képességének becslésére történtek már kísérletek. Az emberi agy hozzávetőleges kapacitása ugyanis mintegy 1100 milliárd ($= 10^{11}$) neuron, amit megszorozhatunk az egyes neuronokból kiinduló kb. 1000 ($= 10^3$) kapcsolattal: ennek alapján több mint 10^{14} kapcsolat van az agyban, és ez még megszorozható kapcsolatonként mintegy 10 bájt szinapszisztípustól függő információval. Ez a kapacitás tehát kb. 10^{15} bájt, azaz petabájt nagyságrendű, ami 1000 terrabájt, vagyis 1000 trillió bit. Ennek a számmisztikaszerű számolásnak önmagában természetesen nincs sok értelme, de azon mindenképp érdemes elgondolkozni, hogy már a nyolcvanas évek legelterjedtebb személyi számítógépe, a Commodore 64 memóriájában (!) is volt 64 kilobájt, ami ugye kicsit kevesebb, mint 64 000 bájt, ami meg valamivel több, mint 500 000 bit, tehát a C64-nek több mint $2^{500\,000}$ lehetséges állapota volt. Ez a szám iszonyú nagy: $2^{500\,000} = 2^{10 \cdot 50\,000} \approx 10^{3 \cdot 50\,000} = 10^{150\,000}$, ami elmondhatatlan mértékben meghaladja a fent kapott 10^{15} -öt. Mégis, mindezek a felfoghatatlanul nagy számok nem jelentenek az emberével semmilyen szempontból sem összehasonlítható képességet: a C64 teljesítményét a rajta futó egyszerű BASIC interpreter, továbbá néhány kezdetleges szövegszerkesztő, és egy-két, mai szemmel még grafikusnak is alig mondható játékprogram jellemezte. A különbség tehát az emberi agy és a személyi számítógép között elsősorban a szervezettségben, azaz a tárhely szervezésében, és nem feltétlenül a méretében van. Mindez annyira igaz, hogy Cuthbertson (2014)

arról számol be, hogy egy japán–német együttműködésben létrehozott szuperszámítógéppel megpróbálták az agy idegrendszerét modellálni, de a 705 024 processzormagos és 1,4 millió gigabájt memóriájú gép az emberi agynak még csak 1%-át képes szimulálni. Jelenlegi hardvereink teljesítménye tehát erre elég, bár a technológiai fejlődés valószínűleg egy újabb robbanás előtt áll, ha hamarosan megjelennek a kvantumszámítógépek.

Ha kilépünk a numerikus közelítések világából, elgondolkozhatunk az egyébként erősen „nyelvi jellegű” Turing-teszten is, hiszen ez az intelligencia fogalmának egy korai közelítéseként jelent meg. Mint ismeretes, ebben a tesztben szerepel egy kérdező és két tesztalany, melyek közül az egyik ember, a másik pedig gép (Turing 1950). A kérdező nem látja, nem hallja a tesztalanyokat: azok írásban reagálnak a feltett kérdésekre, és írásbeli reakcióikkal mindketten megpróbálják meggyőzni a kérdezőt arról, hogy ők gondolkodó emberek. Ha a kérdező ötperces faggatás után sem tudja egyértelműen megállapítani, hogy a két alany közül melyik a gép, akkor a gép sikerrel teljesítette a tesztet. Persze, azóta már sokan kritizálták ezt az elgondolást, mondván, hogy attól még lehet intelligens egy gép, hogy nem képes az itt elvárt emberi módon kommunikálni; és az is igaz, hogy talán az emberek közül se teljesítené mindenki sikerrel a Turing-tesztet.

Az emberi intelligencia definíciói a problémamegoldásra irányuló szellemi tevékenységek összességét próbálják megfogni. Ennek a komplex képességnek részét ugyanúgy képezi a gondolkodás gyorsasága, mint az összefüggések átlátására való képesség, de a memória, sőt sokszor az általános ismeretek is, melyek viszont már kortól, kultúrától, műveltségi szinttől és még rengeteg mindentől függhetnek. Az intelligenciáról talán azt ki lehet jelenteni, hogy egy igen összetett emberi képességről van szó, melynek létezik például logikai, nyelvi, vizuális, zenei, mozgási és még nagyon sokféle megvalósulása. Az intelligenciatesztek is általában ilyen típusú rész-képességeket mérnek. Ezek mérésére különböző okos emberek különféle feladatokat találtak már ki (Neisser és mtai 1996), de egyik sem bizonyult „a” megoldásnak. Azt viszont már tudjuk, hogy ezek a képességek nem füg-

genek az agy fiziológiai méreteitől, ugyanis a kutatók eddig még nem találtak szignifikáns kapcsolatot az agy mérete és az intelligencia között (Wicket és mtsai 1994).

2. Mi is az a gépi intelligencia?

A gépi intelligencia fogalma a (soha pontosan nem definiált) emberi intelligencia analógiájára született. Ezáltal azok nehézségeit is magában hordozza, hiszen a gépi intelligencia azokat a képességeket volna hivatott összefoglalni, melyeket a számítógépek akkor használnak, amikor komplex feladatokat próbálnak megoldani. Shirai és Tsujii (1982) szerint a gépi intelligencia nem más, mint az a készség, mely a számítógépeket alkalmassá teszi az emberi intelligenciával megoldható feladatok ellátására. Persze itt is, mint sokszor, az emberi intelligencia ismét csak nincs pontosan definiálva. Kurzweil (1990) szerint a gépi intelligencia az olyan funkciót teljesítő gépi rendszerek létrehozásának a művészete, amikhez intelligencia szükséges, ha azt emberek teszik. Rich és Knight (1991) – talán kicsit szarkasztikusan – úgy gondolják, hogy a gépi intelligencia megpróbál a számítógéppel olyan dolgokat művelni, amiben pillanatnyilag az emberek egyértelműen jobbak. Dagli (1994) szerint viszont a gépi intelligencia emulálja, vagyis lemásolja az ember ingerfeldolgozási és a döntéshozó képességeit, így ezeknek a rendszereknek autonóm tanulási képességekkel kell bírniuk, és alkalmazkodniuk kell tudni bizonytalan vagy részlegesen ismert környezetekhez. Mivel manapság a tanulás fogalma kulcsszereplővé lépett elő a mesterségesintelligencia-kutatásokban, a gépi tanulásról még szót fogunk ejteni.

3. A generatív nyelvészet és a „gépi intelligencia”

A nyelvészeti kutatások elég korán, a múlt század ötvenes éveinek elején látszottak már érintkezni a számítógépekkel, bár maga

a mesterséges intelligencia fogalma akkor még nem volt ismert, hiszen annak kialakulását egy 1956-os tudóstalálkozóhoz szokás kötni (McCarthy és mtsai 1955). Az a levezetési rendszer ugyanis, amit értelemszerűen a formális logikától vett át a modern nyelvészet, tulajdonképpen egy olyan absztrakt gépet feltételezett, amelyik ezt a levezetést végrehajtja. Tehát az elemzés, amikor a szinkron nyelvészet megjelent, és a strukturalizmus kialakult, olyan fogalommal vált, aminek a napi szintű megvalósításába akkor csak azért nem tudott bekapcsolódni a számítógép, mert még nem létezett. Viszont az a leírási mód, ahogy a levezetés fogalma kialakult, teljesen alkalmas volt arra, hogy a végrehajtást egy intelligens eszközre bízva, ha majd lesz ilyen. Ennek ellenére az ezt követő ötven évben megvalósított generatív grammatika – bármennyire modernnek tűnik is sokak számára – gyakorlatilag seholy nem jár karöltve a modern nyelvi technológiákkal és a mesterséges intelligenciával. Bár a Turing, Post és Thue nyomán Chomsky által – még a transzformációs generatív grammatika előtt – kidolgozott formális nyelvek és az ezek felismerésére képes absztrakt automata (Chomsky 1956) azt sugallták, hogy a modern nyelvészet az ötvenes évek végén a gépi intelligenciák felhasználásának irányába mozdul el, nem így történt. A mesterséges intelligencia valójában sohasem találkozott a generatív nyelvészettel.

4. A mesterséges intelligencia első találkozásai a nyelvészettel

Az ötvenes évek közepén tehát megjelent a mesterséges intelligencia fogalma az informatika világában. Ekkor még mindenki hitt abban, hogy a számítógép az emberi agy egyfajta modellje, hiszen ez idő tájt az emberi intelligencia közelítése látszott az egyik legfontosabb célnak. Ez az idők folyamán változott (bár soha senki nem mondta ezt ki). A legelső intelligensnek tűnő számítógépes produktum, amely a nyelvtudománnyal találkozott, a gépi fordítás volt. Ebben az időben még javában égett a hidegháborús igények által komolyan táplált gépi fordítási láz. Az

automatikus fordítás egy olyan mechanikusnak látszó tevékenységet próbált megfogni, mely a világháborúban gyakran alkalmazott sifírozás-desifírozás mintájára képzelte el az intelligens gépek által elvégzendő munkát (Warren 1949). Az alapgondolat az volt, hogy ha volna olyan gép, ami egy ilyen tevékenységet el tudna végezni, akkor ez a gépi intelligencia egyfajta megnyilvánulásának számítana. Volt persze olyan gondolat is abban az időben, hogy a fordító a valóságban nem is a fordítandó mondat nyelvi elemzésén keresztül oldja meg a problémát, de amikor az első gépi fordítás publikusan megjelent,¹ a sajtó azonnal „elektronikus agynak” keresztelte el a fordító gépet, mondván, hogy ilyesmit létrehozni csak az intelligencia képes – bármi legyen is az (Hutchins 1997). Mivel azonban korábban még nem volt olyan eszköz, ami az emberen kívül ilyesmit csinált volna, kézenfekvőnek tűnt a hasonlat, ami bár nagyon félrevezető, mind a mai napig ott szerepel az újságírók kedvenc fordulatai közt.²

Az ezt követő időszakban voltak azután nyelvészeti közeli mesterségesintelligencia-kísérletek is, mint például az intelligens dialóguspartnert szimuláló ELIZA (Weizenbaum 1966), vagy az egy szűk geometriai világban nyelvi parancsokkal irányítható SHRDLU (Winograd 1972). Ezek az alkalmazások a legtöbb mesterségesintelligencia-könyv bevezető részében ott szerepelnek, de nyelvtudományi szempontból sosem látszottak szignifikáns eredménynek.

5. A statisztikai közelítés „intelligenciája”

A számítógépek kapacitásának növekedésével egyre több szöveg vált tárolhatóvá és kereshetővé. Ez jelentős lépés volt a korpusz-nyelvészeti kialakulásához. A nagyméretű szövegtörzsek ke-

¹ A georgetowni IBM-kutatócsoport munkája eredményeképpen 1954-ben (IBM Press Release 1954).

² A téma jobb kifejtéséhez érdemes elolvasni Neumann János *A számítógép és az agy* című művét, amiben ilyen direkt párhuzam sehol sem szerepel.

zelésére egyre érdekesebb statisztikai módszerek jöttek divatba. Ezek – szakmai kifejezéssel – nagy fedésű (azaz rengeteg, szemmel sokszor nehezen azonosítható nyelvi jelenséget észreévő), de relatíve kis pontosságú (azaz téves észrevételeket is tartalmazó) módszerek. A fedés annak a megfogalmazása, hogy a lehetséges bemenetek hány százalékára tud reagálni egy konkrét algoritmus. A statisztikai rendszerek – szemben a szabályalapúakkal, melyek csak arra reagálnak, amikre van szabályuk – mindig adnak választ, a fedésük nagy. A pontosság annak a mérőszáma, hogy az adott válaszokból hány százalék adekvát. A szabályalapúak, ha van megfelelő szabályuk a bemenet kezelésére, relatíve használhatóbb válaszokat adnak, a statisztikaiak viszont általában nem. Így a statisztikai rendszerek pontossága alacsony. Azt a tulajdonságot, hogy a statisztikai módszerek minden bejövő jelsorozatra tudnak valamit mondani, így arra is, amire közvetlenül nem készítették fel őket, egyesek egyfajta gépi intelligencia megvalósulásának látták. Jelentős különbségek jelentek meg ugyanis a korábbi nyelvészeti felfogáshoz képest: az adatvezérelt közelítés szemben állt az addigi elméletvezérelt közelítésekkel. A statisztikai nyelvészeti módszerek és a generatív nyelvészet viszonya kemény ellentétpár. Érdekesség, hogy a korpusznyelvészetnek van pozitív hozadéka, a „serendipity principle”, és egy negatív hozadéka, a „sparse data problem”. Az első azt jelenti, hogy a tudományos felfedezés sokszor szinte véletlenül történik meg, hiszen a nagy mennyiségű adat szemrevételezésekor eddig nem ismert összefüggések hirtelen felismerése minden előzetes erre való készülés nélkül történhet meg. A másik probléma viszont épp arra utal, hogy mindig lesznek olyan jelenségek, melyek megfelelő működésének kimutatásához nem elegendő az aktuális méretű adathalmaz. Ha a hiányt hasonló típusú, de az eredetivel nem teljesen azonos részadatbázisból vett mintákkal pótoljuk, a rendszer ezt egyfajta „zaj” bevezetésének éli meg, azaz ahelyett, hogy az apróbb részletek pontosabb árnyalásának tekintené, még sokszor azt is elrontja, amit eddig jól „ismert”. Így a statisztikai megoldások kapcsán felmerülhet a kérdés, hogy lehet-e intelligensnek nevezni azt a rendszert, amelyik adott eset-

ben még a szóazonosságot sem ismeri? A statisztikai rendszerek, így a statisztikai gépi fordítás is ilyen, hiszen az nem szavakat, hanem úgynevezett n -gramokat, azaz n darab egymást követő szóból álló, átfedő szócsoportokat³ fordít, és ha valami nem fordult elő a tanítókörpuszában, akkor azt csak mechanikusan tudja „összerakni”, a részeinek a statisztikai alapján, ami nagyon sokszor torz fordításhoz vezet. Általában ezek azok a félrefordítások, amit a felhasználók meglehetősen döbbenetesen fogadnak.

6. Neurális hálók, vektoros reprezentációk, mélytanulás

A 2010-es években megjelenő mesterségesintelligencia-technológiai és -módszertani változások a nyelvi jelenségek gépi kezelésére is nagy hatással voltak. Ezekben az években megjelentek ugyanis a folytonos vektorteres reprezentációk, előtérbe tört a mélytanulás, azaz a neurális hálók megjelenése, illetve egészen pontosan „újrafelfedezése” a legismertebb jelenség. Persze ahhoz, hogy ezek a gépi megoldások működőképeseek legyenek, kellett egy-két algoritmikus megoldás, az elsősorban ma a grafikus processzorok (GPU) segítségével létrejött, a korábbiaknál lényegesen hatékonyabb számítási teljesítmény és rengeteg adat. Ez utóbbi, azaz az „adatéhség” – tudományágtól függetlenül – napjaink legjelentősebb kutatási területein megjelenik. A neurális hálók egyébként már korábban is léteztek, csak egyszerűbb szerkezettel: a bemeneti szint és a kimenet között mindössze egyetlen, a rendszer nevét adó neuronszerű elemekből álló, és ez által neurálisnak nevezett réteggel. A mélytanulásban a bemenet és a kimenet között egynél több neurális szint helyezkedik el. A ma már sok gyakorlati alkalmazásban is megjelenő arcképfeldolgozás (leegyszerűsített) példája jól illusztrálja ezt: az ún. konvolúciós hálóban első szinten az orra, a fültre, a szemre

³ Egy ilyen modellben egy k szóból álló mondatban $k - n + 1$ darab n -gram készül. Így például egy 10 szóból álló mondatban 8 darab trigram van: az 1.–2.–3., a 2.–3.–4., ... és a 8.–9.–10. szavakból álló.

vagy a szájra jellemző görbék, vonaldarabok azonosítása történik meg, majd ezek súlyozott összegéből a következő réteg már „össze tudja állítani” magukat az említett arc-összetevőket, majd egy ezt követő réteg ezekből az arcot. A nyelv esetében sincs nagyon másképp: a magasabb rendű egységek neurális felismerése az alacsonyabb rendűek felismerésével indul. Felmerül a kérdés, hogy ezek a rendszerek mennyiben nevezhetők intelligensnek? Mivel az intelligenciának a korábbi fejezetekben vázolt bizonytalan definíciói nem kedveznek az elméleti közelítéseknek, segítségül hívjuk a gyakorlati működést. A disztribúciós szemantika „újrafelfedezése” elsősorban a neurális hálók első generációjának hardverlehetőségeihez viszonyítva óriási távlatokat nyitott. A disztribúciós szószemantika működtetéséhez szükséges folytonos vektorterjes jelentésrepresentációt úgy kell elképzelni, hogy a bemeneti és a kimeneti szint közötti egyetlen neurális réteg minden egyes bemeneti elemhez (= szóhoz) hozzárendel egy vektort, melynek elemszámát (= dimenzióját) kísérleti alapon a hálózatot építő szakember határozza meg. Ezek, az így létrejött vektorok akkor hasonlóak, ha az általuk reprezentált szavak szókörnyezetei általában nagy hasonlóságot mutatnak, azaz ha a használati szabályaik, tehát jelentéseik (Wittgenstein 1989) hasonlóak. A neurális hálók tehát új szintre emelték az évtizedek óta ismert disztribúciós szemantikai modellek gyakorlati használhatóságát. Egészen pontosan: ezek a reprezentációk a metaforikus értelmezésből a valóság területére helyezték át a szókörnyezet által definiált szójelentést. Sőt, ami váratlan jelenség: a formai és a jelentésbeli szempontok mindegyikének jelenléte mellett a stilisztika is megjelenik, egyfajta további „finomosztályzásként”.

7. A magyar szókészlet disztribúciós szemantikai vizsgálata: az első ígéretes eredmények

A következőkben felvillantjuk egy ilyen szóalapú vektorterjes reprezentáció néhány érdekes, talán elgondolkasztató tulajdonságát. A példák az MTA–PPKE Magyar Nyelvtechnológiai Kutatócso-

port két kutatójának, Siklósi Borbálának és Novák Attilának az eredményeiből valók (Siklósi–Novák 2016). Az ő szövektoros kísérleteik olyan eredményeket hoztak a magyar nyelv esetében is, amik további vizsgálatokat inspiráltak (Novák–Novák 2017). Arra a kérdésre azonban, hogy valójában miket „kódolunk” a szövegeinkben, nem tudunk válaszolni, pedig a használt algoritmusok igen meggyőző eredményeket hoznak. Az első kísérletek magyar szóalakok jelentéseinek hasonlóságát vizsgálták elemzetlen korpuszon (1. ábra). Az ábrákon a 0-val jelzett sorban a lekérdezett alak áll, alatta pedig a hozzá a rendszer által leghasonlóbbnak „gondolt” alakok közül az első nyolc. A két számoszlop közül az első a hasonlóság mértéke (minél közelebb van az 1-hez, annál hasonlóbb), illetve a szó gyakorisága a korpuszban.

0	kenyerek	1	2270
1	zsemlelék	0.8105	283
2	péksütemények	0.8048	997
3	kekszek	0.7972	1046
4	pékárúk	0.7957	771
5	tészták	0.7881	2466
6	lepények	0.7849	202
7	kiflik	0.7843	349
8	kalácsok	0.7841	277

1. ábra. A *kenyerek* szóhoz hasonló szóalakok

Tövesített korpuszon⁴ a fogalmi rokonságok jobban megragadhatóak, hiszen az egyes szavak töveinek gyakorisága egészen más, mint egyes toldalékolt alakjainak a gyakorisága (2. ábra).

⁴ A tövesített korpuszban a szavak helyett azok szótári alakjai és a toldalékok által megfogalmazott nyelvészeti információ állnak. Például: a *baglyok* szó helyett az áll, hogy: *bagoly* PL (azaz: többes szám).

0	franciakulcs	1	255
1	feszítővas	0.8590	846
2	csavarkulcs	0.8445	473
3	csípőfogó	0.8242	345
4	pajszer	0.8219	567
5	hidegvágó	0.8054	156
6	csavarhúzó	0.7984	4369
7	csőfogó	0.7890	111
8	villáskulcs	0.7890	764

2. ábra. A *franciakulcs* szóhoz hasonló szótövek

Amit igazán fontos megjegyezni, az az, hogy a tipikusan nem szótári elemek, tehát a(z akár nem magyar eredetű) tulajdonnevek (3. ábra), a divatkifejezések, szlengek (4. ábra), illetve a félregélt alakok (5. ábra) ugyanolyan biztonsággal jellemezhetők ezzel a módszerrel, mint a normához közelebb álló társaik.

0	Smith	1	30236
1	Wilson	0.8491	19408
2	Thompson	0.8476	6506
3	Harris	0.8447	8168
4	Walker	0.8429	6358
5	Fisher	0.8418	4496
6	Adams	0.8413	7335
7	Taylor	0.8407	20785
8	Wright	0.8360	6385

3. ábra. A *Smith* névhez hasonló szótövek

0	bealszik	1	4325
1	elszundit	0.8061	781
2	elálmosodik	0.7889	2118
3	elbóbiskol	0.7833	1507
4	visszaalszik	0.7742	5858
5	elalszik	0.7646	62217
6	kipurcan	0.7399	1187
7	szunyál	0.7315	974
8	elbambul	0.7285	1953

4. ábra. A *bealszik* igéhez hasonló szótövek

0	rövidnac	1	43
1	pizs	0.7731	180
2	napszemcs	0.7584	37
3	sap	0.7460	374
4	zacs	0.7259	170
5	szemcs	0.7209	37
6	pih	0.7198	149
7	suzuk	0.6943	131
8	nemtomm	0.6795	47

5. ábra. A félregévelt *rövidnac* alakhoz hasonló szótövek

Érdekes észrevennünk, hogy tartalmi és a formai hasonlóság érdekes kapcsolatával állunk itt szemben. Azt már az eddigiekből is érezhetjük, hogy tartalmi hasonlóságra számíthatunk a modell működésekor, de hogy ez finom stilisztikai információk „megéreztetésével” is kiegészülhet, arra először nem is gondolt senki. A (6. ábra) nevei mind magyar női nevek, de a hasonlóság az alapszóhoz egyfajta „stilisztikai érzékről” is tanúskodik.

0	Katalin	1	88546
1	Zsuzsanna	0.8893	30461
2	Ilona	0.8783	33342
3	Ágnes	0.8750	69813
4	Gabriella	0.8735	27494
5	Judit	0.8730	74435
6	Szilvia	0.8483	18932
7	Ildikó	0.8465	55454
8	Klára	0.8442	22176

0	Eufrozina	1	254
1	Jolánta	0.7732	307
2	Konstancia	0.7679	275
3	Gertrúd	0.7469	1530
4	Eugénia	0.7418	342
5	Adelheid	0.7410	185
6	Amália	0.7187	1748
7	Sarolt	0.7168	795
8	Gertrud	0.7093	802

0	Kincső	1	1242
1	Csenge	0.8689	4680
2	Evelin	0.8662	3497
3	Bianka	0.8620	4242
4	Fanni	0.8465	10955
5	Kitti	0.8452	6544
6	Cintia	0.8387	1194
7	Villő	0.8358	542

6. ábra. A *Katalin*, az *Eufrozina* és a *Kincső* női nevekhez hasonló szótövek

A fenti példák csak annak az illusztrálására szolgálnak, hogy a mesterséges intelligencia soha nem volt ennyire közel a nyelvészeti kutatásokkal való együttműködéshez, mint napjainkban. Természetesen ez még csak a kezdet kezdete, hiszen a szóbeágyazási modelleknél lényegesen komplexebb megoldások látnak folyamatosan napvilágot: ezek azt mutatják, hogy szövegeinkben olyan dolgokat is kódolunk, amikről eddig ennyire explicit tudomásunk nem volt. Talán úgy lehetne megfogalmazni, hogy akár még néhány éve is nehezen lett volna elképzelhető, hogy egy-egy inkább csak metaforikusnak gondolt nyelvészeti állítás-

nak gépi igazolására is sor fog kerülni, például annak, hogy a jel jelentése nem más, mint a jel használati szabálya.

8. Összefoglalás

A kérdés, amit igyekeztünk körüljárni, hogy miként találkozott a gépi intelligencia és a nyelvtudomány az elmúlt évtizedekben, odavezetett, hogy most már inkább az a kérdés, hogy a 21. század elején nem egyfajta gépi intelligencia születésének a határán állunk-e? Ha átfutjuk az utóbbi idők idetartozó eredményeit, akkor azt látjuk, hogy a kialakult rendszerek mögött meghúzódó matematika (legalábbis a szakemberek számára) tökéletesen érthető; az internetről ingyenesen letölthető számítógépes implementációk jól működnek, ám mi, emberek nem tudunk a kiinduló és az eredményadatok közti, az általunk létrehozott algoritmusok által numerikusan megfogalmazott (és a példák által is igazoltan működő) összefüggéseknek értelmezést adni. Ráadásul ezekhez a módszerekhez nemcsak rengeteg adat kell, hanem a hatékony működéshez a rendszer matematikai paramétereinek megfelelő beállítása szükséges. Ám épp az a nehézség, hogy az ember nem képes átlátni az algoritmus futása közbeni állapotokat, így nem tud előre jól kiszámítható paraméterbeállításokkal előállni. Igaz, a gépet is és a programokat is mi hoztuk létre, de még ha lépésről lépésre végigkövetnénk is a működését, arra nem jönnénk rá, hogy melyik az a pont, ahol a rendszer „megértette” az emberi nyelv szavainak az értelmét. Ezért a tudományos kutatásban eddig egyeduralkodó módszert, a monoton pontosítást⁵ egy ma még talán furcsa, intuíción alapú kísérleti tevékenység látszik felváltani. A részleteiben nem kontrollálható rendszerek használata közben az ember a megoldást az előző feladatok gyakorlati tapasztalataiból kialakított intuíciónak hagyja.

⁵ A szabályalapú világ sajátja, hogy az eddigi eredményeket javítjuk azáltal, hogy újabb szabályok megfogalmazásával még pontosabb közelítését tudjuk adni a leírandó jelenségnek.

Nem látunk bele ugyanis a mesterséges neurális háló működésébe, így például nem tudjuk, hogy valójában miként kellene elképzelni azt az adott esetben több száz dimenziót, amelyben a vektoros reprezentációk létrejönnek, és nem ismert a válasz arra a kérdésre sem, hogy lehet-e majd az aktuálisan létrejöttnél (nyelvi szempontból) „intelligensebb” modell. Egy kicsit sántító analógiával azt mondhatnánk, hogy hasonló a helyzet ahhoz, amikor egy vizsgált személynek az álmában történetek jelennek meg, ám a vizsgálatot végző szakember csak az eközben zajló kémiai vagy elektromos folyamatokat tudja mérni: az álmokképek a mért numerikus adatok alapján nem „azonosíthatók”. Napjaink neurális hálós módszerei ebben különböznek a korábbi gépi megoldásoktól, hogy még ha jónak tűnik is a kapott eredmény, a köztes rétegekben megjelenő numerikus értékeknek semmilyen interpretációt nem tudunk adni. A bemutatott példák mögötti szóbeágyazási modellben a dimenziószám 300-ra lett beállítva, ami azt jelenti, hogy a program háromszáz független szempontból osztályozza az egyes szavakat a szöveggörnyezetük alapján. Ha bárkit megkérnénk, hogy mondjon 300 ilyen osztályozási szempontot, nem sikerülne, mert a gyakorlati életben nem találkozunk ennyivel. Amikor például barkochba játékot játszunk, igyekszünk néhány független szempont szerint „betájolni” a kitalálandó szót, mint amilyen a méret, a szín, a forma, az eredet – ám ilyen szempontokból legfeljebb néhány tized tudunk összeszedni, így elképzelésünk sincs a rendszerünk által használt több száz dimenzió jellegéről.

Összegzésképpen arra az alapkérdésre, hogy mennyire önjáró ma a gépi intelligencia, azt válaszolhatjuk, hogy egyre inkább. Ha arra a kérdésre is válaszolnunk kell, hogy miként alakítja át a gépi intelligencia megjelenése napjainkban a humán kutatást, akkor mondhatjuk, hogy valószínűleg meglehetősen. Ám ha valaki a határokat feszegetné, és azt kérdezné, hogy hol válik nehézzé a mesterséges intelligencia kutatási felhasználása, a válasz egyértelműen az, hogy ahol tényleg intuíció kell. A gépi megoldások ugyanis jelenleg csak hatékony másolásra képesek, viszont az ehhez szükséges mintákat a program számára nekünk, embe-

reknek kell szolgáltatóknunk. Jelen ismereteink azt sugallják, hogy az új típusú alkalmazások sora tehát nem a tudományelméletet alakítja át, hanem sokkal inkább csak a kutatási gyakorlatot.

Irodalom

- Chomsky, Noam 1956. Three Models for the Description of Language. *IRE Transactions on Information Theory* 2: 113–124.
- Cuthbertson, Anthony 2014. Supercomputer uses 1.4 million GB of RAM to calculate a single second of human brain activity. <https://www.itproportal.com/2014/01/13/supercomputer-takes-40-mins-calculate-single-second-human-brain-activity/>
- Dagli, Cihan 1994. *Artificial Neural Networks for Intelligent Manufacturing*. London: Chapman & Hall.
- Hutchins, John 1997. From first conception to first demonstration: the nascent years of machine translation, 1947–1954. *Machine Translation* 12 (3): 195–252.
- Kurzweil, Ray 1990. *The Age of Intelligent Machines*. Cambridge: M.I.T. Press.
- McCarthy, John – Minsky, Marvin – Rochester, Nathaniel – Shannon, Claude 1955. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>
- Neisser, Ulrich – Boodoo, Gwyneth – Bouchard, Thomas – Boykin, Wade – Brody, Nathan – Ceci, Stephen – Halpern, Diane – Lochlin, John – Perloff, Robert – Sternberg, Robert – Urbina, Susana 1996. Intelligence: Knowns And Unknowns. *American Psychologist* 51 (2): 77–101.
- Novák Arttila – Novák Borbála 2017. Magyar szóbeágyazási modellek kézi kiértékelése. In Vincze Veronika (szerk.): *A XIV. Magyar Számítógépes Nyelvészeti Konferencia előadásai*. Szeged: Szegedi Tudományegyetem. 67–77.
- Rich, Elaine – Knight, Kevin 1991. *Artificial Intelligence*. New York: McGraw-Hill.
- Shirai, Yoshiaki – Tsujii, Jun-ichi 1982. *Artificial Intelligence Concepts, Techniques and Applications*. New York: Wiley.

- Siklósi Borbála – Novák Attila 2016. Beágyazási modellek alkalmazása lexikai kategorizációs feladatokra. In Tanács Attila – Varga Viktor – Vincze, Veronika (szerk.): *A XII. Magyar Számítógépes Nyelvészeti Konferencia előadásai*. Szeged: Szegedi Tudományegyetem. 3–14.
- Turing, Alan 1950. Computing Machinery and Intelligence. *Mind* 49: 433–460.
- Weaver, Warren 1955. Translation. In Locke, William – Booth, Andrew (szerk.): *Machine Translation of Languages*. Cambridge, MA: MIT Press. 15–23.
- Weizenbaum, Joseph 1966. ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the Association for Computing Machinery* 9: 36–45.
- Wickett, John – Vernon, Philip – Lee, Donald 1994. In Vivo Brain Size, Head Perimeter and Intelligence in a Sample of Healthy Adult Females. *Personality and Individual Differences* 16 (6): 831–838.
- Winograd, Terry 1972. *Understanding Natural Language*. New York: Academic Press.
- Wittgenstein, Ludwig 1989. *Logikai-filozófiai értekezés*. Budapest: Akadémiai Kiadó.

Palkó Gábor

A digitális bölcsészet kultúrtechnikái. Virtuális kutatókörnyezetek

Tudatlanságunk a technika kapcsán mérhetetlen.

Bruno Latour

Michel de Certeau a történetírás történetét elemző híres könyvében, melynek talán legfontosabb fejezete *A történeti művelet* címen magyarul is olvasható, a történetírás feltételei között különös hangsúlyt helyez olyan paraméterekre, amelyek, és ez negyven évvel ezelőtt éppen úgy igaz lehetett, mint ma, legtöbbször kívül kerülnek a bölcsészeti kutatás önreflexióján.¹ A „hely” kettős funkcióját értelmezve kifejti, hogy az, egyrészt, lehetővé teszi a kutatások egy adott formáját a közös problémák mentén, míg, másrészt, cenzorként működik, kizár a diszkurzusból más kiindulópontokat. De Certeau érvelésében a történettudomány és a bölcsészeti kutatás gyakran egymást fedő fogalmak, ahogyan azon állítás esetében is, amely a „helyet” a kutatás kereteit meghatározó feltételként írja le, ami összeköti a lehetőséget és a lehetetlent. A konkrét, egyedi hely, ahol a történész vagy általában a bölcsész kutató megszólal, éppen olyan fontos, mint az a hely, amiről a kutatás szól, amiről szó van.² De ha a kutatói megszólalás helye – vagy, mint jelen érvelés esetén, az a hely, amiről szó van – virtuális, beszélhetünk-e egyáltalán fizikai, szociális stb. helyről? Vagy konkrétan: a kutatást meghatározó környezet elemzése közben De Certeau figyelembe veszi-e a digitális környezet lehetőségét? A kérdés persze anakronisztikusnak is tűnhet egy olyan szöveg összefüggésében, amely nemhogy az eScience, de általában a hálózati társadalom (Manuel Castells) kialakulása előtt fogal-

¹ De Certeau 1988.

² De Certeau 1988: 68.

mazódott meg. Éppen ezért meglepő az a tény, hogy a számítógép szerepét de Certeau milyen nagyra értékeli a bölcsészeti kutatás egészére nézve. Amikor azt a gyakorlatot elemzi, hogy a kutatás reflexiós teréből az „egyetem” miként utasítja ki a nem interpretatív praxisokat a „segédtudománnyá” nyilvánítás gesztusával, akkor a felsorolás a felirattan, a kodikológia, a paleográfia stb. után a számítógép-tudományt is a nem reflektált, de reflektálandó gyakorlatok közt említi, kiemelt helyen méghozzá.³

A szerző még egy kontextusban tárgyalja a számítógépet mint a kutatást alapjaiban átalakító jelenséget az idézett esszében. Eszerint csak az a kutatás tekinthető tudományosnak, amely saját „forrásait”, a kutatás „dokumentumait” másképpen alkotja meg, mint ahogyan azok korábban adottak voltak. Új relevanciára hivatkozva a bölcsészeti mű eszközöket, recepteket, dalokat, a termőföldek elhelyezkedését vagy a város topográfiáját dokumentummá formálja, megváltoztatva azok funkcióját. A kutatás, ha tudományos, újraosztja a teret (megint kettős értelemben: az alany és a tárgy terét) intézményesítő cselekvés és átalakító technikák révén. Ez utóbbiak leírására használja De Certeau az apparátus fogalmát, amely egyszerre feltétele, eszköze és eredménye a tudományos kutatást jellemző elmozdításnak vagy újrendezésnek. Az apparátus De Certeau által kiemelt szegmentuma az archívum, amely mint intézményesült technológia meghatározza, lehetséges-e új válaszokat adni másfajta kérdésekre. Vagy másképpen: egy új típusú apparátus teszi csak lehetővé új kérdések és válaszok megfogalmazását. A kutatást alapjaiban meghatározó változásra az archívumi (archiváló) aktivitásban De Certeau példája (François Furet nyomán) a számítógép: a digitális adathordozó mint archívum egy sor változást hoz, terében a vizsgálatnak csak az lehet tárgya, ami – még a programozási folyamatot megelőzően – formalizálható. Egy fontos lábjegyzetben, ahol a számítógép működése és a hagyományos könyvtár között húz párhuzamot az esszé, megjelenik az információs tér fogalma, vi-

³ De Certeau 1988: 69.

lágossá téve, hogy a kutatás helye egészen új értelmet nyer az információtudomány terében, „input” és „output” között.⁴

De Certeau tehát, összességében, nem pusztán azt jósolja, hogy a számítógép részévé válik a tudományos kutatás mindennapi praxisainak, hanem azt, hogy mint technológia új apparátust képez, ami meghatározza a kutatás helyét és a kutatható helyeket egyaránt: azt, hogy miről és hogyan lehet beszélni a humán tudományban. Mindez talán a De Certeau által – ironikusan – segédtudományként aposztrofált praxisok esetében a legegyszerűbb. Az írástörténet, a kézirat, a könyvtörténet olyan területek, amelyeknek a napi gyakorlatait és összefüggésrendszerét a digitális közegbe való áthelyezés igen radikálisan írja újra az elmúlt évtizedekben. A filológia apparátusának digitalizálódása – és ennek széles körű metafilológiai reflexiója – kitűnő példája mindennek. Bernard Cerquiglini *A variáns dicséretében*, amelyet joggal neveznek az „új filológia” alapszövegének,⁵ a (kutatási) tér megváltozásának, amikor is a „könyvszerű elrendezés” átadja a helyét a „képernyőszerűnek”, a kutatás egészét érintő jelentőséget tulajdonít. A digitális apparátus radikálisan átrendezi a kutatás tereit, még hozzá abban a kettős értelemben, ahogy arra De Certeau is utal: a kutatást reprezentáló dokumentum és a kutatás tárgyát képező dokumentum esetében egyaránt. Sőt Cerquiglini ennél is tovább megy, és az áthelyeződést a tinta valóságából az anyagtalanság elektronikus impulzusok világába – Michel Foucault nyomán – a szerzőfunkció megváltozásával hozza összefüggésbe.

„Amit a képernyő olvashatóvá tesz, az mindig a pillanatnyi megragadás, a különböző és mozgásban levő szövegterek, már nem az elhasznált tinta tapintható valóságából, hanem néhány elektromos impulzus anyagtalanságából felépülő szövegek múlékony vizualizációja. A szöveg kétdimenziós stabilitása, melyet egy eredeztető és uraló szubjektumnak tulajdonítottak, már nem jelenik meg az írás ezen új technológiájának alapjánál.”⁶

⁴ De Certeau 1988: 107.

⁵ Déri et al. 2011: 39.

⁶ Cerquiglini 2011: 296–297.

Természetesen nem De Certeau és nem is Cerquiglini az egyetlen, aki az apparátus digitális technológiáinak ilyen radikális szerepet tulajdonít a bölcsészeti kutatás és, általánosabban, a kultúra működésében. Az archívumok elméletei, ezen belül a magát médiaarcheológiának aposztrofáló kutatási irány vagy csoport nagy figyelmet fordít az archívumok digitális áthelyeződésére, az adatterek apparátusának megértésére, különös hangsúlyt helyezve e megértés, a médium átláthatóságának a korlátaira, az archívum „árnyékos oldalára”. (Azt, hogy Wolfgang Ernst vagy Jussi Parikka elméleti írásai vajon mennyiben tárják fel az intézményes digitális archívumok működéséből következő korlátokat és lehetőségeket, amelyeket az apparátus nyújt, egy másik tanulmányban próbáltam meg körüljárni.⁷) Ha nem is konkrétan a számítógép, de általában a kutatás és az apparátus kölcsönviszonyában De Certeau is utal arra, hogy ez az összjáték nem feltétlenül átlátható. Az apparátus meghatározza, hogy hol helyezkednek el a kutatás vakfoltjai, de hogy a vakfoltokra miként és honnan nyílik rápillantás, az esszében nem válik világossá.

Ahhoz, hogy ezt a kérdést értelmezzük a kutatás mai, virtuális környezeteinek összefüggésében, két másik szerzőhöz fordulunk. De Certeau szerint a kutatás maga teremti meg saját dokumentumait, annak az apparátusnak a terében, amelyet használ, és amelyet a tudós át is alakít. A dokumentum megformálásának gyakorlatai a digitális környezetben megváltoznak, de korántsem bizonyos, hogy rendelkezésünkre áll az a nyelv vagy perspektíva, amely ezt a változást képes beláthatóvá tenni. (Itt újra utalhatunk a dokumentum kettős értelmére: a kutatások retorikája szerint valamit dokumentáló forrás és a közlés folyamatai révén keletkező dokumentum kettősségére.⁸)

⁷ Vö. Palkó 2017.

⁸ De Certeau odáig megy, hogy a különféle múltak nem pusztán azok dokumentumaiban maradnak fenn, hanem a történeti műben mint „archívumban” (De Certeau 1988: 46).

Cornelia Vismann *Akten. Medientechnik und Recht* című munkája, amely angolul legalább olyan gyakorta citált kötet, mint németül, a fordításban olyan főcímet kapott, és ez a jelen érvelésben különös jelentőségre tesz szert, amelynek mediatechnológiai kétértelműsége korántsem idegen a könyv érvelésétől: „Files”. Annak magyarázataként, hogy az irattárak bürokratikus rendszerének történetét miért érdemes technikai részletekbe mennően tanulmányozni, Vismann egyrészt a papír nélküli iroda ígéretét, a számítógép fájljainak és a papír aktáknak a kurrens ellentétét, másrészt a szabad hozzáférés mai társadalmi igényét nevezi meg. Ugyanakkor magában a könyvben, amelynek egy részlete egyébként magyarul is olvasható a Helikon folyóirat *Az archívumok elméletei* című számában,⁹ az elektronikus „akták” csak a rövid zárófejezetben kerülnek szóba. Az aktákból a számítógép képernyőjén látható ikonok lesznek, erre utal a fejezet címe is: *Aktákból ikonok*. Vismannt idézve: „A számítógépes felület terminológiáját úgy tervezték, hogy a képernyő előtt ülő felhasználót az akták ismerős [családias] világára emlékeztesse. A menü fül olyan lehetőségeket kínál föl, mint »lista«, »formátum«, »tezaurusz«, »táblázat«, és az olyan utasítások, mint a »másolás«, »törlés«, »mentés« a felhasználót virtuális kancelláriává vagy kancellárrá változtatják. Azáltal, hogy magába sűrít egy teljes adminisztratív irodát, a számítógép a bürokrácia alapszabályát valósítja meg, amely szerint az adminisztráció technikáit átviszik az államtól az egyénhez: a kora újkori kancelláriák sajátos kormányzati gyakorlatait a köz stílusára, az abszolutista adminisztráció központjait az egyén íróasztalára, a védelmi minisztérium első nagyszámítógépeiről az otthoni asztalra.”¹⁰

Ha Vismann-nak igaza van, és a számítógép „felszíne”, a felhasználói felület az analóg aktakezelés ismerős világát hivatott felidézni, az a nyelv pedig, amelyen megszólít bennünket, a digitális műveleteket analóg gyakorlatokkal írja le, akkor az a különbség, ami a fájlok analóg és digitális kezelését, az analóg és

⁹ Vismann 2014.

¹⁰ Vismann 2008: 163.

digitális archívumot elválasztja egymástól, nehezen lesz felszínre hozható. Az, hogy a Microsoft Windows ablakain keresztül egy virtuális irodai „Asztralra” látunk rá, ismerős közegét képezheti a számítógéppel való interakciónak, de eközben az, hogy mi történik a mélyben, a motorháztető (vagy akár: az asztal?) alatt, az rejtve marad.

Ezt a jelenséget, vagyis a szakadást a képernyőmédiium és a számítógép mint digitális apparátus működése között, hogy egy némiképp eltérő érvrendszert idézzünk fel, Niklas Luhmann a felület és a mélység közötti ellentét segítségével írja le a *Die Gesellschaft der Gesellschaft* médiatörténeti áttekintésében.¹¹ A gép kapcsolatban áll ugyan a képernyő felületén keresztül a felhasználóval, ugyanakkor homályban marad.

„A felszín most a képernyő az emberi értelem rendkívül behatárolt igénybevételével, a mélység ellenben a láthatatlan gép, amely manapság pillanatról pillanatra, magától újrakonstruálja önmagát, például amikor reagál a használatra. A felszín és a mélység közötti kapcsolat létrehozható a parancsok által, amelyekkel a gépet vezérlik, hogy a képernyőn vagy a nyomtatás révén valamit láthatóvá tegyen. Ő maga azonban láthatatlan marad.”¹²

A fekete doboz, és ezt nem az elektronikus médiiumok fenti leírása, sokkal inkább a kettős kontingencia fogalma kapcsán válik egyértelművé,¹³ Luhmann-nál sohasem nyitható fel, a számítógép zárt belső világa, akár a pszichés rendszer, mindig megközelíthetetlen marad. Ha a kutatás digitális környezeteire innen tekintünk, kevés esély marad az apparátus értelmezésére.

Bruno Latour számára az emberi és a nem emberi közötti interakció mindig ki van téve egy a luhmannihoz hasonló elrejtettségnek vagy átlátszatlanságnak. Latour kifejezése, a black box látszólag közel áll ahhoz a láthatatlansághoz, ami a számítógép önmagát pillanatról pillanatra újrakonstruáló működéseit rejti el Luhmann-nál. Valójában, és ez a bölcsészettudomány

¹¹ Luhmann 1998.

¹² Luhmann 1998: 304.

¹³ Krause 2005: 17.

és a számítógép viszonyában döntő különbség, Latour számára nem lehetetlen felnyitni a fekete doboz tetejét, hogy láthatóvá váljék, mi van a motorháztető alatt. Példája az írásvetítő, ami, amíg működik, alig is létezik, amint elromlik, még egyes részei is létezővé válnak, megannyi fekete dobozzá.¹⁴ A zárt fekete doboz funkcionális részekre bomlik, amelyek külön-külön szintén fekete dobozként működhetnek, de egymáshoz kapcsolódásuk átláthatóvá válhat. Amit ez a bepillantás a fekete dobozba feltár, az nem egységes, tárgyszerű objektum: egy gép működése, hanem olyan emberi és nem emberi cselekvők együttműködése, akik/amelyek közel vagy épp távol vannak egymástól. Ennek a feltárulásnak a jelölője a „technikai” melléknév, és ezzel a technikai apparátus és a kutatás viszonyát érdemben érintő kérdéshez értünk el ismét: „A »technikai« elsősorban egy alprogramra, vagy egymásba illeszkedő alprogramok sorára vonatkozik (...) Amikor azt mondjuk, ez egy technikai pont/pillanat, ez azt jelenti, hogy kitérőt kell tenni, el kell térni egy pillanatra a fő feladattól és újra kell kezdeni a cselekvések megszokott sorát, ami egyedül figyelmet érdemel. Egy fekete doboz kinyílik, majd hamarosan újra bezárul, láthatatlan lesz ismét a cselekvések fő sodrában.”¹⁵

A latouri érvelésben ennek a feltárulásnak a feltétele a technikai működés fennakadása, ami az eredeti célok megghiúsulását is jelentheti, azt, hogy egyszer csak valahol egészen máshol találjuk magunkat, mint ahová készültünk: az apparátus működésének hibája annak átalakulásával járhat. „»[E]lőször egy technikai problémát kell megoldani«. Itt a kitérő nem biztos, hogy visszavisz bennünket a főútra, mint a fenti esetben, de veszélybe sodorhatja az eredeti célt. A technikai többé nem pusztán kitérő, hanem gát, úttorlasz, egy kerülő, egy hosszú áthelyeződés kezdete...”¹⁶

¹⁴ Latour 1999: 183.

¹⁵ Latour 1999: 191.

¹⁶ Latour 1999: 191.

A technikai ebben az értelemben a tudományosság előfeltételeként azonosított elmozdulás, áthelyeződés ágensévé válhat a De Certeau-i terminológia szerint.

A tanulmány eddigi részében Michel de Certeau, Cornelia Vismann és Bruno Latour egy-egy belátását idéztük fel, arra való tekintettel, vajon a bölcsészeti kutatás és a digitális környezet összefüggésében mennyiben tűnnek felhasználhatónak. Vállalva a durva leegyszerűsítés kockázatát, így foglalhatnánk össze a fenti kontextusokat.

De Certeau érvelése felől láthatóvá vált, hogy a kutatás „helye”, az apparátus meghatározza, egyáltalán milyen kutatói kérdések fogalmazódhatnak meg, vagyis hogy a kutatói környezet digitális átalakulása radikálisan átírja a kutatási problémák és az azokra adható válaszok körét, miközben a tudományos kutatás maga is az apparátus újrendezésére kell hogy vállalkozzon.

Vismann akták és fájlok között húzott párhuzama arra világított rá, hogy óvatosan kell bánnunk az ismerős, sőt családias felületekkel, amelyek az analóg iroda világot imitálják, miközben digitális technikákat hajtanak végre, melyek maguk is hatalomtechnikai gyakorlatok áthelyeződései.

És végül Latour érvelése a technikairól felvillantotta annak lehetőségét, hogy a komplex technikák megakadása képezte szünetben emberi és nem emberi, időben/térben közeli és távoli összjátékát pillantsuk meg a fekete doboz felnyitásakor, még hozzá olyan egymáshoz kapcsolódó alprogramok együttműködésésként, ami kizárja a tisztán emberi kontroll lehetőségét, az emberi uralmát a nem emberi felett.

Reális kérdés, ezek az elméleti belátások vajon hogyan használhatóak a bölcsészeti kutatás konkrét digitális gyakorlatainak összefüggésében. Ennek bemutatására két konkrét példát említenék. Az egyik a digitális filológia területéről származik, a másik a távoli olvasás (*distant reading*, Franco Moretti) köréből.

A Petőfi Irodalmi Múzeum és az MTA BTK Irodalomtudományi Intézet a DigiPhil projekt (digiphil.hu) keretei között vállalkozott az Arany János-levelezés kritikai kiadás öt nyomtatott kötetének digitális átírására és kutathatóvá tételére. A mun-

kában részt vevő szerkesztők számára irodai szoftverkörnyezetet imitáló felhasználói felület készült. Egy ilyen felület nem enged bepillantást a fekete dobozba, a kötet jelölőnyelvi kódolását végző szerkesztő nem látja az ismerős/családias felület mögött azt a kódot és bonyolult munkamenetet, amelynek specifikációja digitális filológusok és Arany-szakértők együttműködésével jött létre. Latouri értelemben itt a cselekvők nem egyidejű együttműködésének példáját azonosíthatjuk, ahol a kutatás helyének, a digitális környezet egymáshoz kapcsolódó fekete dobozainak a kreatív megalkotása, mint De Certeau-i értelemben véve tudományos kutatótevékenység, illetve a környezeten belül folyó munka elválik egymástól.

Ha – mintegy megakasztva a munkafolyamatot, a dobozfedelelet felemelve – rápillantunk a kutatókörnyezet mint apparátus működésére, valóban emberi és nem emberi, közeli és távoli, múltbéli és jelenkori páratlanul bonyolult összjátékát látjuk magunk előtt. Az idősíkok első nyalábja kézenfekvő: Arany levelezésének 19. századi postai útja a megírástól a feladáson át a megérkezésig, illetve a nyomtatott kritikai kiadás elkészítésének évtizedes folyamata (amely eleve számol a levelek évszázados kulturális kontextusával, a recepcióval is) a 20. század második felében. A digitális közegbe való átkerülésnek magának is komplex az időbelisége. Az első fázisban a digitális filológusok és irodalomtörténészek megalkotják azt a modellt, amellyel a mintegy háromezer levél adathálózattá alakítható, ezen belül döntenek a konkrét jelölőnyelvi kódolás felépítéséről, az adatok és a metaadatok lejegyzésének és kapcsolódásának módjáról (miközben korábban és más szakértők által megállapított jelölőnyelvi ajánlásokat [például TEI XML], metaadatszabványokat [például MARC] és jó gyakorlatokat követnek). Ezután a kódolás specifikációjának ismeretében más digitális bölcsész szakemberek egy adott szoftverkörnyezetben (Oxygen XML Editor) létrehoznak egy olyan felhasználói felületet, amely más, ismert (irodai) szoftverek felületét imitálja (gombok elnevezése és elhelyezése, menüszerkezet), de valójában a korábban megállapított – és a felület használója számára láthatatlan – kódolási folyamatot hajt végre

File Edit Find Project Options Tools Document Window Help

Fejlec Bekezdés Vers Versszak Verssor Szerk. Jegyzet Lábjegyzet Más szerző jegyzete Idézet a levélről Címzés Cím Verscím Alcím Személyn

Törlés Javítás Beszúrás Margó Szerk. beszúrás Hiány Sic Eredeti Feloldott Emendálás Bizonytalan Aláhúzás Dőlt Alsóindex Felsőindex Egybeírva Elválasztv

Untitled1.xml x

level_17_18_19 meta levelMegjelenes

MENTOVICH FERENC – ARANY JÁNOSNAK **Név:** Marosvásárhely **Marosvásárhely** **P. sic.** Mikor:

Kedves **Id:** OHA:96679 **Név:** Arany János **János** om!

Vasánap délelőtt van, szép tavaszi nap, milyen még az idén nem volt – s én még sem megyek a szabadba sétálni, hanem nekem szemrehányásokat tesz és atkozottul stichíroz; **P. stichíroz** – bökdös (né.) azon mulasztást, mely abból áll: hogy könnyelműsége hajlandó természetem félnek: hogy eddigi mulasztásomból részvétlenséget, megfélelkezést és hidegséget gondolnom, s nekem valódi ünnepe, midőn levél vagy a kőszá hír utján rólad és kedves családotról hallok valamit; elhiszed barátságod, s azt el nem veszítené az egész életre kiérdemelt legőszintébb oajtásom egyike. Ne engedje az ég édes **It** körülmények úgy hozni magokkal hogy arra szükség lenne – én az első közé tartoznám, kik érted és kedves családó kifejténi. Ha híven elődbe tárhatnám azon kedélyhangulatot, melylyel e sorokat leírtam: bizonyára meggyőzne téged szavaimból semmit sem von le az, hogy ismeretségnk után oly késő idő múlva történik, ez nem az új bor felpesdülése, **P. sic.**

Annyi mondani és kérdeni valóm van, hogy nem tudom hol kezdjem. Mindenek előtt hallj valamit állapotunkról. Én egész férjnek, mint ha a nőnek novekszik **P. sic.** a hasa. A Maros partja mellett **Név:** Erdély bérci közt ismét év megszakasztott volt. Olyan az én életemben ezen hat esztendő mint egy rozsz és sületlen regényben egy gyönyörű és szép mely alkalommal irigység tárgya vagyok, drága pénzen vennék meg tőlem hogy hallottam **Id:** OHA:96679 **Név:** Arany János **OHA:96679 Név:** Arany János **AJ** kedves nótáját (**P. Kodály-Gyulai 1952. 4 38. 4**) **P. Kodály-Gyulai 1952. 4 38. 4** kigördülni; hogy ferblizt értéket éri el (né.) **Id:** JHA:142388 **Név:** Tompa Mihály **Tompá** val és **Id:** JHA:114359 **Név:** Jókai Mór **J** Losonczy László **Losonczi** t az **olájet** az **olájet** – az **Oláj nótá** t (**P. Kodály-Gyulai 1952. 4 15. 4**) romjait. Anyagi állapotom tűrhető, nagy előnye az itteni tanári állomásoknak, **P. állomásoknak** – állásoknak **4** hogy kény egy-egy urfi kosztost, sőt magán órakat is lehet adni ur házaknál. Családom is jól van. Feleségemnek úgy látszik használ a nőnek, mindjárt nagy leányaim lesznek. **Id:** **Név:** Mentovich Gyula **Gyula** szépen halad a hegedűben és ta

Hát ti hogy vagytok? bajlods-e most is fészédülésekkel? A kedves komámasszony meg van-é régi jó egészségében? **M** **Név:** Losonczy László **Losonci** val, viccelődjünk **Id:** JHA:126347 **Név:** Szilágyi Sándor **Szeleki** vel?! **P. A bal**

A záró címke előtt nem szerepelhet szököz.

Text Grid Author

C:\Users\Fellegisz\Documents\VSL\Projects\Untitled1.xml

[Név](#) [Helynév](#) [Dátum](#) [Bibliográfia](#) [Bibl. rövidítés](#) [AJ levél utalás](#) [Táblázat](#) [Új sor](#) [Új cella](#) [HELP](#)
[rtva](#) [Elkülönítve](#) [Összehúzza](#) [Lehúzza](#) [Tollhiba](#) [Elkenődött/Tintafojt](#) [Olvashatatlan betű](#) [Olvashatatlan rész](#) [+Meta sor](#)

[Után:](#) [Előtt:](#) [▶](#) **Április 18-án 1858.** [◀](#)

nem leültem asztalom mellé hogy jóvá [▶](#) [sic](#) [◀](#) tegyek egy mulasztást, mely mián az önvád darab idő ota nyugodni nem hágy, mely
 hogy neked még nem írtam, pedig már mindjárt két éve hogy körötökből eljöttem. Ha nem ismernéd halogató s egy kis
 ígert olvasandsz ki - így azonban remélem elhiszed szavamra: hogy gondolatban sokszor fölkerestelek és kereslek föl, jól esik reád
 ed ha azt mondom: hogy [Név:](#) Magyarország [▶](#) Magyarország [◀](#) ből magammal hozott legszentebb emlékeim közé tartozik a te
[Id:](#) OHA:96679 [Név:](#) Arany János [▶](#) János [◀](#) om! hogy semmi tekintetben barátid segélyére szorúl - de ha a sors a
 dodért tesznek valamit, itt nem volnék hanyag, a halogató [Id:](#) JHA:119411 [Név:](#) Mentovich Ferenc [▶](#) Mentovich [◀](#) itt erélyt fogna
 ram öszintességéről. Bocsáss meg ez érzékenykedésért - de ládd nem volt még alkalmam érzésem öszintességét neked bevallani,
 l de az óbor olajos bouquet-jának, [▶](#) bouquet-jának [▶](#) - zamatának (fr.) [◀](#) illatos zamatjának tekintheted azt.

szség dolgában jól vagyok, azt mondják házom, hasam nő, mire én azt szoktam felelni: hogy családod háznál sokkal jobb ha a
 ét folytatását élem azon gyermek- és ifjú [▶](#) [ko](#) [▶](#) [kori](#) [▶](#) napjaimnak, melyet a [Név:](#) Magyarország [▶](#) Magyarhon [◀](#) ban töltött hat
 zépen kikerekített epizód, sokszor beszéltek köztetek eltöltött eleményeimből [▶](#) [sic](#) [▶](#) társaságban itteni barátim- és ismerőseimnek,
[▶](#) Arany János [◀](#) komoly bajusza alol mázsás élceket, s „a vasvári verebeket” [▶](#) „a vasvári verebeket” [▶](#) - [Id:](#)
[▶](#) [▶](#) ferbliztem [▶](#) - ferbli: magyar kártyajáték, amelyben az nyer, aki négy ugyanazon színből való lapjával a legnagyobb
[▶](#) Jokai [▶](#) val [▶](#) [sic](#) [▶](#) - s hogy (ezt olvasd [Id:](#) OHA:96681 [Név:](#) Arany László [▶](#) Laci [▶](#) nak) hogy láttam [Id:](#) [▶](#) [Név:](#)
[▶](#) [▶](#) [▶](#) eltáncolni - de én mindezeket nem adnám el - de örzöm őket, mint a szerelmes a kapott rozsa vagy zöld levél száradó
 ényelmes szállásokkal vannak ellátva, fizetésem - mióta a directorságot rám erőltették - fölmegy 800 e forintra, tarthat az ember
 l az erdélyi levegő - kivált mióta a mult nyáron fürdön volt jobb erőben van, szeretném a nyáron is elküldetni. gyermekeim [▶](#) [sic](#) [▶](#)
 tanulásban is.

Megérem - é még vagy egyszer hogy az általa sütött híres turos rétes mellett a jó miskolciti szörpölgetve komázunk [Id:](#) [▶](#)
 al lapszálon e sor mellett (feltehetőleg [Id:](#) OHA:96681 [Név:](#) Arany László [▶](#) Arany László [▶](#) tól származó) kérdőjel. [▶](#) [▶](#) [▶](#) [Id:](#)

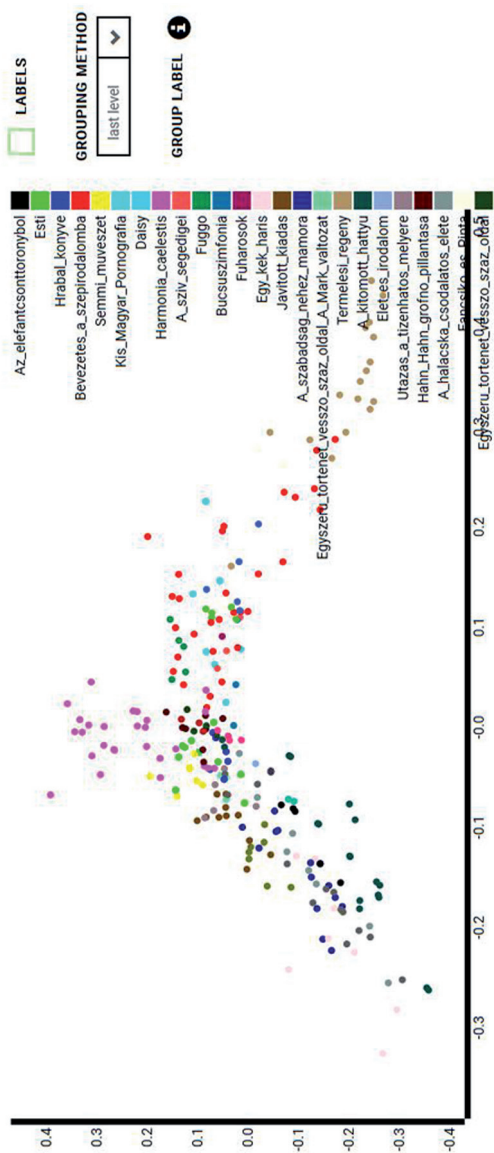
U+0000	Modified
--------	----------

1. ábra. Arany János egy levele a DigiPhil virtuális kutatókörnyezetében

a specifikációnak megfelelően. A felületet manipuláló (használó) szerkesztő számára az a folyamat is rejtve marad, hogy az általa készített kódolt szövegen milyen ellenőrző algoritmusok futnak, amelyek figyelmeztetnek a tipikus hibákra, szintaktikai szabálytalanságokra (például dátum rossz formátuma, hiányzó kötelező metaadat). A levelezés és a kritikai jegyzetek szövegében a szerkesztő jelöli a nevet, a felület (máskor és mások által összegyűjtött névlista alapján) javaslatot tesz a személynevek azonosítására, mely javaslat alapján a szerkesztő döntést hoz az azonosításról, majd – gondosan dokumentálva a kódon végrehajtott minden egyes változást – a rendszer egy szerverre továbbítja a kódolt dokumentumot. Ennek további feldolgozása, ellenőrzése és közzététele már egy másik környezetben történik.

A második példa egy olyan stilometriai szolgáltatás, amelynek tesztelése során a Wrocław Műszaki és Tudományegyetem, a Digitális Irodalmi Akadémia és az ELTE Digitális Bölcsészeti Központ működik együtt. A szolgáltatás, a Multilingual WebSty célja természetes nyelvi elemző algoritmusok segítségével feldolgozott szövegek hasonlósági vizsgálata. A szolgáltatás, szemben a felhasználói interakciót kizáró, teljesen automatizált megoldásokkal, lehetetlenné teszi, hogy figyelmen kívül hagyjuk egymásba ágyazott alprogramok sora működik együtt a szöveghasonlósági statisztikák és ezek vizualizációja létrehozásában. A szolgáltatás nem elrejt, hanem nagyon is előtérbe állítja azt a technikai apparátust, amely emberi és nem emberi, közeli és távoli cselekvők interakcióját foglalja magában, és amely paraméterek sorával befolyásolható, így olyan eredményeket produkál, amelyek mindig hangsúlyosan relatív érvényűek, kutatói értelmezést feltételeznek, és újraértelmezést provokálnak.

A szolgáltatás tesztelése során Esterházy Péter prózai szövegeit elemeztük a Digitális Irodalmi Akadémia által biztosított szövegtörzskorpuszon. A szövegstatistikai kvantifikációt a szövegek morfológiai elemzése előzte meg: a kapott eredmények számítógépes nyelvészeti kutatások eredményeképp létrejövő – különféle hatékonyságú – külső eszközökre épülnek. (Érdekes tény, hogy a számítógépes morfológiai elemzés hibái nem feltétlenül csökkentik



2. ábra. Esterházy Péter prózájának szöveghasonlósági vizsgálata – a „Text similarity analysis system” (WebSty) vizualizációja

radikálisan a szöveghasonlósági vizsgálatok, például a szerzőazonosítás eredményességét.) A szerzői korpuszt az elemzés során a szolgáltatás véletlenszerűen egyenlő szövegszegmentumokra bontotta, és azokat – a nyelvi elemzés eredményeit felhasználó statisztikai módszerrel – csoportosította. A rendszer alapbeállítások mellett is látványos eredményt hozott, az azonos művekből származó szegmentumokat jó hatásfokkal rendelte egymáshoz, míg, magasabb absztrakciós szinten, az életmű nyelvi-stilisztikai tendenciáiról hírt adó, irodalomtörténeti értelmezéseket alátámasztó (vagy épp cáfoló) vizualizációkat produkált.

A jelen érvelés szempontjából legalább annyira érdekes, hogy egy ilyen komplex stilometriai szolgáltatás – mint latouri értelemben vett fekete doboz – nem értelmezhető heterogén, térben, időben, kompetenciákban egymástól radikálisan különböző, emberi és nem emberi „cselekvők” együttműködésének feltételezése nélkül. Kezdve az elemző szoftvereket nyelvi korpuszon „betanító” számítógépes nyelvészről a matematikai modelleket alkalmazó statisztikuson át a szövegeket kódoló, illetve értelmező irodalmárig, nem beszélve az alkotóiktól (térben és időben) elváló számítógépes algoritmusokig.

A heterogén technológiai apparátus elemzése nélkül a mind szélesebb körben alkalmazott virtuális kutatókörnyezetek elemzése, reflexiója lehetetlen, ugyanakkor a széttartó kompetenciákat újfajta egységbe foglaló apparátusra való rálátás képessége egyre kevésbé adott a magányos kutató egyéni horizontján.

Irodalom

- Cerquiglioni, Bernard 2011. A variáns dicsérete. A filológia kritikai története. In Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel (szerk.): *Metafilológia* 1. Budapest: Ráció Kiadó. 219–297.
- De Certeau, Michel 1988. *The Writing of History*. New York: Columbia University Press. Magyarul: Michel de Certeau: *A történeti művelet*. In Benda Gyula – Szekeres András (szerk.): *Az Annales*. Budapest: L'Harmattan. 2007.

- Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel 2011. Előszó.
In Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel (szerk.):
Metafilológia 1. Budapest: Ráció Kiadó. 9–44.
- Krause, Detlef 2005. *Luhmann-Lexikon*. Stuttgart: Lucius & Lucius.
- Latour, Bruno 1999. A Collective of Humans and Nonhumans: Following Daedalus's Labyrinth. In *Pandora's Hope. Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.
- Luhmann, Niklas 1998. *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel (szerk.) 2011. *Metafilológia* 1. Budapest: Ráció Kiadó.
- Palkó, Gábor 2017. Media archaeology of institutional archives? *Studia UBB Digitalia* 62 (LXII) June, Issue 1: 75–82.
- Vismann, Cornelia 2008. *Files, Law and Media Technology*. Stanford, CA: Stanford University Press.
- Vismann, Cornelia 2014. Az irodától az adatvédelemig. *Helikon Irodalomtudományi Szemle* 3: 365–399.

Debreczeni Attila

Az elektronikus kiadás színe és visszája

1. Bevezetés

Az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport immár hosszabb ideje foglalkozik elektronikus kritikai kiadások fejlesztésével, jelenleg tizenegy önálló kiadásból áll közös adatbázisra épülő portálrendszerünk.¹ A kiadás koncepciója, textológiai és technikai megoldásai hosszú idő alatt formálódtak ki, s ez a folyamat – elektronikus médiumról lévén szó – természetesen most sem tekinthető lezártnak. Számos tanulság adódik eddigi munkánkból, most azonban nem ezek és nem a megtett út ismertetése a cél.² A jelen dolgozat problémák exponálására vállalkozik, mert az elektronikus kritikai kiadás nemcsak vonzó és távlatos lehetőségeket testesít meg, de válaszra váró elvi és gyakorlati kérdések sora merül fel vele kapcsolatban. A válaszok megtalálásának monográfiákat szenteltek már, tengernyi a szakirodalom, a jelen keretek között még csak összegző áttekintésüket sem lehet célul kitűzni. Egyetlen kérdéskör felvázolását

¹ Ezek a következők: életművek (Csokonai, Kazinczy, Szerdahelyi György Alajos, Gyöngyössi János), levelezések (Magyar Írók Levelezése, Ráday család, Berzsenyi), folyóiratok (*Orpheus*, *Magyar Museum*) elektronikus kiadásai, egy tematikus dokumentum-adatbázis a magyar nyelv-újításról és Kazinczy jegyzőfüzeteinek adatbázisa. A kiadások a 'deba.unideb.hu/deba' portálazonosítóval ellátott címen címen érhetőek el (például deba.unideb.hu/deba/levelezes), a Társportálok menüpontban mindegyik kiadás szerepel, ahogy a kutatócsoporti honlapon is (textologia.unideb.hu). A jelen tanulmány az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport programja keretében jött létre.

² Erre nézve lásd korábbi tanulmányomat: Debreczeni 2014.

választottam témául: mi a különbség a nyomtatott és az elektronikus tudományos kiadás között? vagyis miben áll és mivel jár a mediális váltás e területen? milyen új kutatási potenciált teremt és milyen eddig nem létezett kérdéseket vet fel mindez?

Felvetődhet a gondolat, hogy az emez alapkérdések körüli vizsgálódás talán már nem ígér sok újdonságot. Ám ha csak a digitális tudományos kiadások elméletéről és gyakorlatáról tavaly megjelent gyűjteményes kötet³ tanulmányaival szembesülünk is, rögtön láthatóvá válik, mennyi még az eldöntetlenség. Ahogy Hans Walter Gabler az előszóban fogalmaz: „Az előttünk álló feladat ezért a tudományos kiadást mint digitális tudományos kiadást megvalósítani. Ez megköveteli a médium sajátosságainak a maga teljességében való feltárását. Jelenleg a digitális közegben megvalósult kiadások még nagymértékben a nyomtatott kiadások imitációi maradnak. Amit azonban a mediális váltás megkövetel, az a teljes újrakonceptualizálás és következőképpen újramodellezés.”⁴ A digitális tudományos kiadásnak⁵ sem megvalósulásai nincsenek nagyszámban, sem konszenzuális teóriája, minden számottevő eredmény ellenére sem.

2. A mediális különeműség kritériumai

A kérdéskör vázolásához kiindulásként az elektronikus kiadás egy hazai definíciós kísérletét választottam, annál is inkább, mert ez egyfajta szakmai-hivatalos szentesítéssel rendelkezik. Az MTA

³ Driscoll–Pierazzo 2016.

⁴ „The task ahead is therefore to realise the scholarly edition as a digital scholarly edition. This demands exploring the medium’s potentials to their full extent. At present, editions realised in a digital environment still tend to remain largely imitative of scholarly editions in print. What the medial shift requires however is a thorough re-conception, and in consequence re-modelling” (Gabler 2016: XIV).

⁵ A ’digitális tudományos kiadás’ a ’digital scholarly edition’ fordításaként használatos a jelen tanulmányban, a magyar terminológiában inkább az ennek megfelelő ’elektronikus kritikai kiadás’ fogalma a gyakoribb.

Textológiai Munkabizottsága 2004-ben tette közzé az irodalmi szövegek tudományos kiadásainak alapelveit rögzítő dokumentumát,⁶ s benne külön kiadástípusként szerepel az elektronikus kiadás, ugyanolyan leírási szerkezetben, mint a többi, régről ismert kiadástípus. Ebben az időszakban én voltam a bizottság elnöke, a társelnöke Kecskeméti Gábor, közös munkával készítettük el a szövegjavaslatot, majd alakítottuk folyamatosan a különböző testületi viták és döntések alapján, egészen addig, míg azt az Akadémia I. osztálya elfogadta. E dokumentum vonatkozó része azzal a megjegyzéssel zárul, hogy az elektronikus médiumban zajló gyors változások az alapelvek rendszeres felülvizsgálatát teszik szükségessé. Ilyenre eddig nem került sor, indokoltnak látszik tehát, hogy a mintegy másfél évtizedes meghatározásunkat szemügyre vegyem az újabb teoretikus megközelítések és az azóta megszerzett kiadási tapasztalatok fényében.

Az *Alapelvek* a következő meghatározását adja az elektronikus kiadásnak: „Egy mű vagy egy életmű összes szövegforrását feltáró és teljességében rögzítő kiadás, amely lehetőséget teremt a változatok együttes olvasására, a közöttük való szabad átjárásra.” E meghatározás kulcsa a szövegfelfogás, amely lényegében egyezik a genetikus kiadások meghatározásában szereplő szövegfelfogással. Mindkettő szakít a főszöveg egyeduralmával és a változatok, források egyenrangúságát vallja. A különbség közöttük a változatok, források megjelenítésében és kezelésében mutatkozik, vagyis elsősorban technikai jellegű, így *önmagában* nem elégséges a mediális különeműség meghatározására. Létezhetnek és léteznek is ugyanezen szövegfelfogáson alapuló papírkiadások, még ha a nyomtatott könyv médiuma nem is adekvát ehhez a szövegfogalomhoz, mert a teljesség elérése és a források együttes olvasása, a közöttük való átjárás nehezen biztosítható, illetve alig használhatóan bonyolult. De nem elvileg lehetetlen,

⁶ Alapelvek az irodalmi szövegek tudományos kiadásához. *Irodalomtörténet* 2004. 328–330. Elérhető a Textológiai Munkabizottság honlapjáról is (<http://textologia.iti.mta.hu/alapelvek.pdf>).

tehát ez a szövegfelfogás szükséges, de nem elégséges feltétele egy elektronikus kiadásnak.

Ugyanez mondható el a digitális médiumról is *önmagában* véve: azért, mert egy kiadás képernyőn olvasható, még nem lesz elektronikus kiadás. A digitális és a digitalizált ezen megkülönböztetése követi az *Alapelvek*ben az idézett alapdefiníciót,⁷ s Kecskeméti Gábor kísérőtanulmányában a különbségtétel tartalmaként azt írja, hogy az „elektronikus kritikai kiadásnak el sem képzelhető papírváltozata: létmódja, filozófiája, szövegkezelő elve kizárólag elektronikus környezet számára megalkotott”.⁸ A tudományos digitális szövegkiadások teoretikusa, Patrick Sahle pedig úgy fogalmaz a már idézett tanulmánykötetben, hogy az elektronikus kiadás nem adható ki nyomtatásban tartalmának és funkcióinak lényegi vesztesége nélkül,⁹ majd megadja definícióját: „A tudományos digitális kiadások tudományos kiadások, melyek elméletükben, módszerükben és gyakorlatukban a digitális paradigma által meghatározottak.”¹⁰ A kérdés igazából az lehet, hogy meg tudjuk-e nevezni a digitális paradigma meghatározó vonásait, amelyek egy tudományos kiadást digitális tudományos kiadássá tesznek. Számos kiváló kísérlet született már erre, de mint a Gabler-előszó jelzi, a munka még nincs elvégezve. S ha konszenzus övezte, zárt, definitív rendszert remélünk ettől, nem is igen lesz elvégezve, mert ahogy a nyomtatott kiadások típusainál, úgy ebben az esetben is életidegen lenne az uniformizálásra törekvő előírásrendszer.

Az *Alapelvek* ebben a szellemben minimumkövetelmények meghatározására törekedett minden kiadástípusnál. Az elektro-

⁷ „(Ennek megfelelően nem tekintjük önálló kiadástípusnak a digitalizált kiadásokat: ezek a papírkidásoktól csak a számítógépes adathordozóban különböznek, nem pedig szövegfelfogásukban.)”

⁸ Kecskeméti 2004: 322.

⁹ „A digital edition cannot be given in print without significant loss of content and functionality” (Sahle, 2016: 27).

¹⁰ „Scholarly digital editions are scholarly editions that are guided by a digital paradigm in their theory, method and practice” (Sahle 2016: 28).

nikus kiadás esetében a mediális váltás szempontjából a legmeghatározóbbak a szövegközlés alcím alatt megfogalmazott kritériumok: „A szövegváltozatok betűhűek (lehetőleg hasonmásban is szerepelnek), ezek mellett olvasószöveget is közölni kell, amely már tartalmazhat emendálásokat. A szövegeket a Unicode kód-tábla szerinti karakter- és egy alkalmas jelölőnyelv szerinti struktúra-kódolásban kell rögzíteni, vagy lehetővé kell tenni, hogy ilyenné konvertálva legyenek kinyerhetők.” E két mondat valójában a digitális paradigma azon két jellemzőjét írja le, melyeket Patrick Sahle a *szövegrepresentáció* és a *transzmedializáció* fogalmaival nevez meg.¹¹

A szöveg három reprezentációja, azaz a kézírathú átírás vagy kritikai szöveg (*critical text*), az emendált, megtisztított olvasószöveg (*reading text*) és a digitális faksimile napjainkban már az elektronikus kritikai kiadások meghatározó és megszokott elemei. Tara L. Andrews egyenesen úgy fogalmaz, hogy ha másban nem is, de ebben konszenzus van a digitális kiadás meghatározása körül folyó diskurzusban.¹² Noha nem elképzelhetetlen egy kis szövegkorpust e hármasság szerint nyomtatott kiadásban közölni, de ezek dinamikus egymás mellé rendezése, ami ugyancsak elvárás az elektronikus kiadásban, lényegében már megoldhatatlan.

Az *Alapelvek*, mint láttuk, óvatosan és megengedően fogalmaz a jelölőnyelvi kódolás előírása tekintetében, részben nem írja elő azt explicite, részben nem nevez meg kötelezően használandó jelölőnyelvet. Napjainkra ez az óvatosság okafogyottá vált, hiszen a jelölőnyelvi kódolás, a szöveg köré és a szövegbe épített metaadatok a digitális tudományos kiadás egyik legfőbb ismérvévé, a mediális különbözőség kritériumává váltak. Eppen ezért nevezi meg Patrick Sahle transzmedializációként a jelölőnyelvi kódolást, amely a kiadást alkotó szövegekből egyedi információforrásokat hoz létre, melyek a rájuk vonatkozó tudást egységben tárolják. Az is eldőlt, hogy az XML vált az általánosan

¹¹ Sahle 2016: 31–32.

¹² Andrews 2013: 62.

használt jelölőnyelvvé, a TEI ajánlásoknak megfelelően. Abban a tekintetben vita folyik, hogy milyen mértékű lehet és legyen a TEI–XML standardizáció. Vannak vélemények, amelyek olyannyira túlzottnak tartják e keretrendszer rugalmasságát, hogy az már a standard működését veszélyezteti, én inkább azokkal tartanék, akik ezt a rugalmasságot a TEI–XML standard lényegének tekintik.¹³ Ahogy a nyomtatott kritikai kiadások számára nem írhatunk elő részletekbe menően egységes és kötelező tartalmi formátumot, úgy szükséges a jelölőnyelvi struktúra részleteit is az anyag sajátosságaihoz szabni, megmaradván persze a TEI–XML standard keretei között, annak belső logikájából építkezve, egy adott kiadáson belül következetesen.

3. Adatbázis-szerkezet és hálózatosság

Az *Alapelvek* az elektronikus kiadás kötelező részei között a médiumra jellemző sajátosságokat is felsorol (megjelenítő program és annak néhány funkciója), de érezhetően itt a legnagyobb a bizonytalanság. E tekintetben más megközelítés célszerűbbnek tűnik ma már, az elektronikus kiadás nem elsősorban a publikációs felület felől értelmezhető. Egy nyomtatott kiadás esetében a kiadás tartalmi lényegét jelentő adat, a kiadás és a publikáció (data–edition–publication)¹⁴ legfeljebb az elkészítés során választható el egymástól, s alkothat időbeli sort, a kinyomtatással ezek zárt egységgé rögzülve produktummá válnak. Az elektronikus kiadás ezzel szemben folytonos keletkezésben van e három elem viszonyrendszerében. Az *adat* a transzmedializáció során információs alapegységgé alakított szöveg, amely a hasonló metanyelvi kódolásban létező szövegekkel együtt adatbázist alkot; a *kiadás* az adatok feldolgozásának szerkezeti és funkcionális kere-

¹³ Lásd erről az idézett tanulmánykötet bevezetőjében és több tanulmányában is (Driscoll–Pierazzo 2016: 12–14, 74–78, 189–200), valamint Andrews 2013: 62–63.

¹⁴ Sahle 2016: 36–37.

te; a *publikáció* ennek megjelenítése html-környezetben (mivel eldőlt a dvd és a hálózati kiadás „versenye” is az utóbbi javára). Az elektronikus kritikai kiadás tehát az adatbázis adataiból a definiált kiadási szerkezetben a felhasználói kérésnek megfelelően *real-time* keletkezik újra és újra az adott publikációs környezetben. Ennek megfelelően külön lehet csak leírni az adatszerkezetét, külön a kiadást strukturáló elemeket, funkciókat és külön a megjelenítési környezetet. De talán a mediális különmeműség kritériumaként elegendő az adatbázis–kiadás–publikáció szétváló hármasságát és ezek folyamatszerű együttműködését megadni.

Az a tény, hogy az elektronikus kritikai kiadás adatbázisra épül, amelyből a felhasználói kéréseknek megfelelően algoritmusok alapján időről időre generálódik, azt is jelenti, hogy nincs állandó lineáris szerkezete, mint a nyomtatott médiumban létező kiadásoknak. Az adatbázisban tárolt szövegek mint információs alapegységek egymással hálózatos viszonyban vannak, amely relációkat a metaadatok rögzítik, jellemzően a TEI–XML file-ok fejlécében. E viszonyokban a szövegek filológiai jellemzői testesülnek meg, s válnak megjeleníthetővé a kiadási szerkezetben, funkciókban kódolt lekérdezési lehetőségek szerint.¹⁵ Az azonban, hogy aktuálisan milyen lineáris rendben jelenik meg egy szöveg, felhasználói döntés eredménye. Épp ezért nem indokolatlan azt állítani, hogy míg a nyomtatott kiadásnak olvasója, addig az elektronikus kiadásnak felhasználója van, s magát e kiadást is joggal nevezhetjük inkább *working place*-nek vagy *research site*-nek.¹⁶ Az így felfogott elektronikus kiadás nem könnyebben használható könyv immár, hanem olyan új kutatási lehetőségek tárháza, amelyek a lineárisan rögzített nyomtatott médiumban elképzelhetetlenek.

¹⁵ Kutatócsoportunk elektronikus kiadásainak e megoldásairól lásd Debreczeni 2014, e megoldások filológiai alapfogalmairól (szövegforrás–szövegállapot–szövegváltozat–szövegidentitás) lásd Debreczeni 2012: 17–21.

¹⁶ Rasmussen 2016.

A szövegek hálózatos szerkezetében megnyíló lehetőségek mellett az elektronikus médiumra tervezett kiadásoknak a másik egészen új karakterisztikumát a szövegek szegmentálása jelenti. A metanyelvi kódolás az adatbázisban egy egységként kezelt szöveg relációinak rögzítése mellett a szöveget fel is bontja, a TEI standard szerint azonosítva annak részeit. Ebből következően ha egy nyomtatott és egy elektronikus kiadásban azonosnak tűnik is egy szövegegység, a metanyelvi kódolás miatt valójában mégis alapvetően különböznek egymástól, s ennek a szövegvizsgálat így megnyíló lehetőségei miatt különös jelentősége van. A szöveg ugyanis a metanyelvi kódolás által különféle gépi vizsgálatokra tehető alkalmassá, de maga a szöveg kódolt szerkezete önmagában is feldolgozás tárgya lehet (metadata harvesting).¹⁷ Épp ezért ezen adatszerkezet közvetlen publikálása egyre általánosabb az elektronikus kiadásokban.¹⁸

A szövegek szegmentálása az alapja a kommentárok készítésének is, s e téren ugyancsak egészen új lehetőségek nyílnak az elektronikus kiadásban. Don Fowles már mintegy két évtizede lelkesen üdvözölte, hogy a kommentár elveszíteni látszik terjedelmi korlátait, és az interaktivitás révén használata is újradefiniálódik.¹⁹ Valójában azonban itt sem pusztán egy könnyebben használható könyvre ismerhetünk: a kommentár műfajának alapkaraktere módosul. A nyomtatott kiadásban a kommentár per definitionem egy másik, elsődlegesnek tekintett szöveghely-

¹⁷ Éppen erre épül kutatócsoportunk együttműködése a Digiphil-projekttel, amely a metadatok feldolgozása és konvertálása segítségével szolgáltatja az adatokat nagy európai projektek (például az Europeana) számára, így a feldolgozott kiadások közvetlenül elérhetővé válnak e gyűjtőrendszerekből is.

¹⁸ Egy pillanatra megállva, hadd emlékezzek meg arról, hogy ezt, az XML file-ok publikálását szorgalmazta Labádi Gergely kutatócsoportunk elektronikus Csokonai-kiadása esetében is, melynek lektora volt. Lektorai jelentése 2017. szeptember 1-én kelt, minden bizonnyal egyik utolsó írása volt ez. Véleménye megerősít minket választott utunkon, köszönettel örizzük.

¹⁹ Fowles 2011: 493–498.

hez van rendelve, még akkor is, ha a médium sajátosságai miatt előállható ismétlődések miatt esetleg több szöveghelyet kiszolgáló jegyzetszótárként is valósulnak meg (hiszen a jegyzetszótárból nem vezet explicitté tett út vissza a szöveghez, legfeljebb csak a mutatókból, vagy ha a jegyzetszótárt a mutatóval keresztezik). Az elektronikus kiadás megoldja az ismétlődés problémáját azáltal, hogy maga a kommentár is az adatbázis önálló rekordjává válik, amely tetszés szerinti helyekhez rendelhető. Minthogy azonban ezáltal az adatbázis ugyanolyan önállósággal bíró része lesz, mint az 'elsődleges szövegek', a viszony per definitionem kétirányúvá válik, azaz lehetőség nyílik az adott kommentárhoz rendelhető szöveghelyek közvetlen lekérdezésére is. Így mondjuk egy szövegben megemlített mű kommentárja nemcsak a mű azonosítását végzi el a szöveghelyhez rendelt (s esetleg a szerző kereshető a névmutatóból), hanem az adott mű összes említése lekérdezhetővé válik a szövegtörzson belül, közvetlenül a kommentárból is.

4. Néhány dilemma

E mediális sajátság a vonzó új lehetőségek megteremtése mellett kérdések sorát veti fel. Egy kommentár alapvető rendeltetése szerint egy másik szöveghely magyarázatát kell hogy megadja. Ez számos esetben 'szöveghely-semleges' információ, hiszen mondjuk Goethe születési és haláladatai mindig ugyanazok, Debrecen mindig ugyanazon geokoordinátákkal azonosítható, és a 'xr.' rövidítés is mindig 'krajcárt' jelent. Az ilyen kommentárok önálló rekordjai bátran hozzárendelhetők a szövegbeli előfordulásokhoz kontextustól függetlenül, legfeljebb az lehet kérdéses, hogy egy Debrecent gyakran említő szövegben vagy a sokszor előforduló 'xr.' rövidítés esetében minden esetben meg kell-e adni a feloldó-azonosító kommentárt. Az elektronikus kiadás adatbázis-logikája azt diktálná, hogy adat és adat között ne legyen különbség, a technikai lehetőség pedig megvan rá. A befogadás ugyanakkor nem igényli, hogy mond-

juk közvetlenül egymás után tízszer ugyanazt az információt el lehessen érni egy-egy felugró ablakban. De ha nem adjuk meg minden előfordulásnál a magyarázatot, akkor milyen gyakran kell és célszerű? Ez egy gyakorlati szöveggondozói probléma, de végső soron az az elvi kérdés rejlik benne, hogy meddig van értelme egy szöveg linearitását felbontani, meddig terjedhet a szövegszegmentálás?

Létezhet azonban Goethe és Debrecen olyan említése, amely kontextuális magyarázatot is kíván a pusztá azonosításon túl, sőt, talán inkább ez tekinthető a kommentár műfajára jellemzőnek. Ebben az esetben éppen az adatbázisjelleg nem tűnik adekvátnak, hiszen az adott információ 'szöveghely-specifikus', leginkább az adott helyhez kell kapcsolni technikailag is, ami betagozza azt a másik szöveg lineáris rendjébe, szöveghely és kommentár sajátos kontaminációját teremtve meg ezáltal.²⁰ Az egymással tartalmilag összefüggésben lévő szöveghely-semleges és szöveghely-specifikus információk persze technikailag összekapcsolhatóak, s szükséges is egy olyan kommentárrendszer megteremtése, mely ezeket egymásra építi. De a kommentár saját műfaji alapkarakterénél fogva ettől még mediálisan némileg zárványszerű marad az adatbázisra épülő elektronikus kiadásban.

Akkor tehát feszültség lenne az új médium és a hagyományos filológia között? Tulajdonképpen logikus fejleménynek tekinthető az a felvetés, hogy az elektronikus kiadások maguk is a digitális médium sajátosságainak megfelelően készüljenek. A tudományos digitális kiadás Patrick Sahle meghatározásában a digitális paradigmát követő tudományos kiadás, amiből következően rendelkeznie kell a tudományos kiadások ismérveivel, azaz textológiai-filológiai értelemben tudományosnak kell lennie. Digitális tudományos kiadás készíthető nyomtatott kiadás adaptálásával, ami a mediális különeműség miatt (mint láttuk) számos nehézséget vet fel még a szövegfogalmában egyező genetikus elvű kiadások esetében is. És digitális tudományos kiadás

²⁰ Fowles 2011: 494.

készíthető eredendően a digitális médiumra tervezve, ebben az esetben az a kérdés merül fel, hogy alkalmazhatóak-e tisztán digitális filológiai módszerek, vagy az eredendően digitális kiadások megalapozásaként sem kerülhető el a hagyományos textológiai-filológiai feldolgozás, minden nehézség ellenére sem.

Tara L. Andrews a digitális filológia szükségessége és létjogosultsága mellett érvel, de kénytelen megállapítani, hogy a kéziratos szövegek automatikus felismerése, átírása, összehasonlításuk, feldolgozásuk jelenleg még nem mutat érdemben használható eredményeket.²¹ Mások, így Matthew James Driscoll és Elena Pierazzo a technikai nehézségekből fakadó kezdetlegesség mellett elvi kérdésként azt is felvetik, hogy mi a szerkesztő szerepe az így felfogott filológiában, a filológus-szerkesztő háttérbe szorítható-e a gépi módszerekkel, illetve hogy azok ténylegesen háttérbe szorítják-e.²² E módszerek távlatait illetően csak óvatosan fogalmazhatunk, hiszen húsz éve sem látszott, ami napjainkra megvalósult, de a belátható jövőre nézve talán kijelenthető, hogy az elektronikus kiadások a hagyományos textológiai-filológiai módszerekkel fognak készülni, miközben az új mediális minőséget az eredendően digitális közegre való tervezés és a jelölőnyelvi kódolás jelenti.

Az elektronikus kiadások esetében nemcsak a hagyományos filológusi szerep kérdőjeleződik meg és értékelődik át, hanem magának a kiadásnak az identitása is. Egy nyomtatott kiadás meghatározott szerkesztői névhez vagy nevekhez kötve, készen és lezártan jelenik meg, s válik egyértelműen azonosítható bibliográfiai adattá. Az elektronikus kiadásnak már a megnevezése is kérdéses, ezek gazdag tárházából elegendő talán most csak a projektet, archívumot, adattárat, gyűjteményt említeni a kiadás mellett.²³ E nevek mindegyike sajátos jellemzőket hangsúlyoz, az azonban közös bennük, hogy nyitottak, azaz folyamatként, nem eredményként értelmezhetőek. Ez természetesen egyfelől üdvöz-

²¹ Andrews 2013.

²² Driscoll–Pierazzo 2016: 6.

²³ Price 2009.

lendő, hiszen a folyamatos korrekció és bővítés nemcsak megoldható, hanem elemi jellemző e kiadásokban. De akkor miben is áll egy ilyen kiadás önazonos volta, mikortól kiadás egy kiadás, illetve a változások által nem lesz-e teljesen más, mint addig? Mire is hivatkozunk egyáltalán ebben az esetben? Természetesen kaphat DOI-számot a kiadás és megadható a letöltés ideje is, ezek azonban éppen az alapsajátosság, a folyamat jelleg kezelésére nem igazán alkalmasak, lévén statikusak és időponthoz kötöttek. Nem arról beszélek tehát, hogy milyen vicces, mikor valamely hivatal, mint egy dokumentumot, úgy kéri pdf-ben egy elektronikus kiadás feltöltését, hanem az elektronikus kiadás elvi különműségéből adódó problémákat jelzem.

Mindez azzal is összefügg, hogy az elektronikus kiadások ma már csoportmunkában készülnek, s a teljesítmény nem feltétlenül elkülönült részekben ölt testet, amelyek így névhez köthetők lennének. Egy kiadás adatbázisában a szöveg hagyományos-textológiai filológiai feldolgozása és kódolása más és más személyhez köthet, amint megint máshoz az informatikai fejlesztés, vagyis a végzett munka nem szövegegységek, hanem feladatkörök szerint különül el. Nyilván megvannak a technikák többszerzős munkaként való elkönyvelésükre, de ez itt némileg mégis más. Mert mi is az egység, amihez a név köthető? A kiadás vagy a kiadás egyes szövegei? Mi a helyzet akkor, ha az adatbázisban elhelyezkedő ugyanazon szöveg több kiadásban is helyet kap (például egy levelezés darabjai a szerző és a címzett szerzői kiadásában)? Egy szövegekódolást végző szakember, aki ma még ritka, mert egyszerre hagyományos és digitális filológiai ismeretekkel bír, hogyan építhet publikációs listát a hagyományos filológiai munkákat is nehezen kezelő tudománymetrikai rendszerünkben? A kérdések hosszan sorolhatóak. Egy részük ugyan „csak” gyakorlati jellegű, de egyéni kutatói pályákat határoz meg, éppen azokét, akikre oly igen szükségünk lenne.

Az elektronikus kritikai kiadás mediális jellemzői tehát vonzó és távlatos kutatási lehetőségeket kínálnak, ugyanakkor megválaszolásra váró elvi és gyakorlati kérdések egész sorát vetik fel.

A tudományos digitális kiadás Joris van Zundert kifejezésével élve egyfajta laboratóriumi környezetet kínál a bölcsészettudományi kutatások és az informatikai tudomány együttműködése számára.²⁴ Jelenleg e laboratóriumban a közös nyelv keresése folyik a két különböző irányból érkezettek között. Ez nem könnyű feladat, és nem is ígérkezik gyorsan megvalósíthatónak. Hogy ismét van Zundert metaforáját idézzem,²⁵ egy kreol nyelv hosszú és bonyolult folyamat során alakul ki (s a digitális bölcsészet esetében most ne firtassuk, ki a gyarmatosító és ki a bennszülött). A pidgin nyelv viszont a maga tört egyszerűségében addig kielégíti a kommunikációs szükséglet elemi igényeit. Úgy tűnik, mi most még mosolygós jó szándékkal, széles gesztusokkal, egy ilyen pidgint beszélünk.

Irodalom

- Andrews, Tara L. 2013. The third way: philology and critical edition in the digital age. *The Journal of the European Society for Textual Scholarship*: 61–76.
- Debreczeni Attila 2012. *Csokonai életművének kronológiai rendje*. Budapest, Debrecen: Akadémiai Kiadó, Debreceni Egyetemi kiadó.
- Debreczeni Attila 2014. Kritikai kiadás papíron és képernyőn. In Czifra Mariann – Szilágyi Márton (szerk.): *Textológia – filológia – értelmezés*. Debrecen: Debreceni Egyetemi Kiadó. 26–39.
- Driscoll, Matthew James – Pierazzo, Elena (eds.) 2016: *Digital Scholarly Editing. Theories and Practices*. Cambridge, UK: Open Book Publishers.
- Fowles, Don 2011. A filológia mint kommentár és a kommentár mint filológia az elektronikus médiumok korában. In Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel (szerk.): *Metafilológia 1*. Budapest: Ráció Kiadó. 476–501.

²⁴ Zundert 2016: 89.

²⁵ Zundert 2016: 87–88.

- Gabler, Hans Walter 2016. Foreword. In Driscoll – Pierazzo (eds.) 2016. XIII–XV.
- Kecskeméti Gábor 2004. A textológiai munka egyes problémáiról – az új textológiai alapelvek közrebocsátásakor. *Irodalomtörténet* 317–327.
- Price, Kenneth M. 2009. Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name? *Digital Humanities Quarterly* 3.3. <http://www.digitalhumanities.org/dhq/vol/3/3/000053/000053.html>
- Rasmussen, Krista Stinne Greve 2016. Reading or Using a Digital Edition? Reader Roles in Scholarly Editions. In Driscoll–Pierazzo (eds.) 2016. 119–133.
- Sahle, Patrick 2016. *What is a Scholarly Digital Edition?* In Driscoll–Pierazzo (eds.) 2016. 19–39.
- Zundert, Joris van 2016. Barely Beyond the Book? Driscoll–Pierazzo (eds.) 2016. 83–106.

Kozák Dániel

Digitális szövegtörzsek a klasszika-filológiában¹

1. Bevezetés

Gregory Crane, a digitális bölcsészet professzora a lipcsei egyetemen² és a klasszika-filológia digitális fordulatának jelenleg talán legismertebb proponense egy 2004-ben megjelent tanulmányában azt írja, hogy a klasszika-filológia és a számítógép történetéről egy ideális világban nem kellene beszélni, mivel a diszciplína művelését segítő digitális eszközök nem eléggé speciálisak ahhoz, hogy indokolt legyen egy „digitális klasszika-filológiát” elkülöníteni az általánosabb értelemben vett „digitális filológiától”, illetve „digitális bölcsészettől”.³ A holt nyelveken írott szövegek tanulmányozása felvet ugyan speciális problémákat is, de valójában minden egyes nyelvnek, élőknek és holtaknak egyaránt megvannak a maga sajátosságai, melyek kezelésére digitális eszközeinket alkalmassá kell tenni. Végző soron pedig a feladat minden esetben nagyon hasonló: a túlnyomórészt nem digitális közegben született szövegek digitális kódolása a lehető legtöbb

¹ A tanulmány az Emberi Erőforrások Minisztériuma ÚNKP-17-4 kódszámú Új Nemzeti Kiválósági Programjának támogatásával készült.

² A Lipcsei Egyetem digitális klasszika-filológiai (és más filológiai) projektjeiről lásd az egyetem Digitális Bölcsészet Központjának honlapját: www.dh.uni-leipzig.de/wo (a jegyzetekben és a bibliográfiában feltüntetett online tartalmak elérési dátuma egységesen 2018.02.25).

³ Crane 2004: 47. Crane tanulmánya Theodore Brunnernek, az alábbiakban tárgyalt TLG-adatbázis egyik alapító szerkesztőjének a klasszika-filológia és számítógép történetét összefoglaló korábbi írására reagál (Brunner 1993).

információ megőrzésével, s az így létrejött digitális szövegek kutathatóvá tétele egyes szövegenként és lehetőleg egybefüggő korpuszként is, minél több nyelvi és tartalmi paraméter alapján.

Két okból mégis indokoltnak látszik külön tanulmányt szentelni a klasszika-filológia digitális eszközeinek. Egyrészt: amint azt Crane maga is hangsúlyozza, a számítógép bölcsész tudományi használatának hajnalán – a kilencvenes évek végéig, vagy akár a kétezres évek közepéig – a klasszika-filológia a rokon diszciplínáktól sok tekintetben eltérő utakon és más ritmusban közeledett a mai értelemben vett digitális bölcsészet felé. Ha tehát elméleti szempontból nem indokolt is, egy történeti áttekintés keretei között mégis hasznos lehet a digitális klasszika-filológia elkülönített tárgyalása. Másrészt: a klasszika-filológia elméletét és módszertanát tárgyaló szakirodalomban mindeddig meglepően csekély mértékben jelent meg az alábbiakban tárgyalt digitális eszközök vizsgálata. Még sok tekintetben feltárára vár és önreflexió tárgyává teendő, hogy használatuk miként és milyen mértékben változtatta és változtatja meg a diszciplína művelését.

Nem vállalkozom arra, hogy a digitális eszközök teljes spektrumát áttekintsem: kifejezetten a digitális szövegkorpuszokat tárgyalom majd, minthogy ezek – nem pedig, példának okáért, a digitális kritikai szövegkiadások – megjelenése gyakorolta a legjelentősebb hatást a klasszika-filológusok munkájára.⁴ Jelen tanulmány első részében a szövegkorpuszok kifejlesztésének történe-

⁴ A digitális klasszika-filológiai (illetve tágabb értelemben: ókortudományi) projektek és eszközök legteljesebb – ám sajnos nem minden tekintetben naprakész – listáját lásd wiki.digitalclassicist.org. Néhány magyar vonatkozású projektet érdemes továbbá kiemelni e helyütt. Ilyen Catullus költeményeinek online kritikai kiadása Kiss Dániel gondozásában (Catullus Online, www.catullusonline.org), Adamik Béla kutatócsoportjának nyelvészeti-epigráfiai adatbázisa (Császárkori latin feliratok számítógépes nyelvtörténeti adatbázisa, lldb.elte.hu), valamint a Szépművészeti Múzeum Antik Gyűjteményének Hyperion projektje, a kutatási eredmények széles, nem kizárólag szakmai közönséghez való eljuttatása szempontjából kiemelten fontos digitális enciklopédiák egyik példája (www2.szepmuveszeti.hu/hyperion).

tét s ezen eszközök mai állapotát tekintem át, második részében pedig néhány szempont szerint azt igyekszem körüljárni, hogy a digitális korpuszok használata miként illeszkedik egyrészt a klasszika-filológia hagyományos, „analóg” módszertanába, s mennyiben alakítja át azt másrészt, különös tekintettel az intertextuális értelmezésre, vagyis annak vizsgálatára, hogy különböző szövegek mennyiben és milyen módon alakítják egymás jelentését és értelmezését olyan esetekben is, amikor a szövegek közti hasonlóságot minden bizonnyal nem valamely szerzői szándék hozza létre.

2. A digitális szövegtörzsek kifejlesztése

A tanulmány alapját képező konferencia-előadásra készülve áttekintettem a Digital Classicist címen, párhuzamosan Londonban és Berlinben megrendezett szemináriumsorozat utolsó három évadában elhangzott 51 előadás címét és absztraktját.⁵ Tanulságosnak tűnik az előadások tematikus megoszlása:

epigráfia, papirologia, kodikológia	9 előadás
3D szkennelés, megjelenítés, nyomtatás; régészet	7
adatok vizualizációja, hálózatelemzés	7
szemantikus web, <i>linked open data</i> technológiák	7
digitális eszközök a (felső)oktatásban; hallgatók bevonása a kutatásba	5
természetes nyelvfeldolgozás (NLP)	5
digitális földrajz, topográfia	5
digitális szakirodalom: hivatkozások kezelése, katalogizálás, publikálás	3
a digitális bölcsészet általános módszertani és kulturális kérdései	2
digitális szövegtörzsek	1

⁵ www.digitalclassicist.org/wip/index.html. Nem vettem figyelembe a 2015–2017-es évadokban elhangzott néhány, az ókortudomány valamely más ágát (például egyiptológia) képviselő előadást.

Nyilvánvalóan lehetséges lett volna más kategorizálás, és figyelembe vehettem volna más előadásokat, publikációkat is; reményeim szerint azonban a lista ebben a formában is alkalmas arra, hogy alátámassza fenti kijelentésemet a digitális szövegtörzsekre irányuló tudományos önreflexió viszonylagos hiányáról. Mindössze két előadó – egyikük éppen a már említett Gregory Crane – tárgyalt a digitális bölcsészettel és annak keretei között a klasszika-filológiával kapcsolatos általános módszertani és kulturális kérdéseket. Ez mégis csupán a második legnépszerűtlenebb témának bizonyult: az utolsó helyre kerültek éppen a tanulmányom tárgyát képező digitális szövegtörzsek, ráadásul a velük foglalkozó egyetlen előadó sem általában tárgyalta ezeket, hanem egy speciális adatbázist: a műkénéi görög nyelvemlékeket összegyűjtő korpuszt mutatott be.

Meglepőnek tűnhet e téma alulreprezentáltsága. Azt sugallja, hogy a klasszika-filológusok az antik szövegeket kifejezetten azok materiális valójában: mint feliratot, papiruszt, kódexet vizsgálják digitális eszközökkel, s újabban a természetes nyelvfeldolgozás (NLP) eszközeit is adaptálni igyekeznek az ógörög és a latin nyelvhez. Mi a helyzet azonban magukkal az antik irodalmi szövegek lehető legnagyobb részét elérhetővé és kereshetővé tévő adatbázisokkal? Nem arról van szó, hogy a klasszika-filológia ne foglalkozna kifejlesztésükkel. Valójában ezek a korpuszok immár mintegy harminc éve léteznek, bár használatuk természetesen csak fokozatosan, a személyi számítógépek és különösen az internet terjedésével párhuzamosan vált a kutatók hétköznapijainak részévé. Az alulreprezentáltságnak az említett előadás-sorozaton tehát inkább az lehet az oka, hogy bizonyos tekintetben a téma mára „lerágott csontnak” tűnhet a klasszika-filológusok szemében: a klasszika-filológusok (és informatikus kollégáik) ma jellemzően más jellegű, informatikai szempontból sok esetben bonyolultabb digitális eszközök kifejlesztésén munkálkodnak, és értelemszerűen ezekről a munkálatokról számolnak be előadásaikban és írott (jellemzően persze online) publikációkban.⁶ Ez

⁶ Lásd például a Heidelbergi Egyetem *Digital Classics Online* open-access folyóiratát (journals.ub.uni-heidelberg.de/index.php/dco).

már önmagában is ezen adatbázisok korai kifejlesztésének negatív mellékhatása lehet – két további mellékhatásra később még visszatérek.

A digitális filológia történetének összefoglalásaiban megke-
rülhetetlen két korai, a latin nyelvű irodalom kutatását segítő
eszköz: Roberto Busa *Index Thomisticus*, Aquinói Szent Tamás
műveinek az IBM közreműködésével már az 1940-es években
fejleszteni kezdett számítógépes konkordanciája,⁷ melynek se-
gítségével a teológus terjedelmes munkásságának egészére vo-
natkozóan lehetett lekérdezni egy-egy szó előfordulásait és azok
közvetlen szöveggörnyezetét, valamint David Packard 1960-as
években létrehozott digitális konkordanciája a római történetíró,
Titus Livius *Ab urbe condita* című részlegesen fennmaradt, de
még így is meglehetősen terjedelmes szöveget alkotó munkájá-
hoz. A több szerző műveit felölelő, általánosabb célú ógörög és
klasszikus latin szöveggörpuszok evolúciójának rövid összefogla-
lását azonban valamivel később érdemes kezdenünk.⁸ 1972-ben
a Kaliforniai Egyetemen egy ógörög lexikográfiai adatbázis kifej-
lesztése ürügyén összegyűlt klasszika-filológusok végül azt a dön-
tést hozták, hogy *Thesaurus Linguae Graecae* (a továbbiakban:
TLG) címmel mégsem lexikográfiai adatbázist, hanem ennél
sokrétűbb felhasználást lehetővé tévő digitális szöveggörpuszt
hoznak létre, melynek tartalmaznia kell az összes fennmaradt
ógörög irodalmi szöveget a homéroszi eposzoktól kezdve a Kr. u.
2. század végéig (ez mintegy 27 millió szó terjedelmű görpuszt
jelent). Az „irodalom” fogalma ez esetben igen tágan értelme-
zendő. Vonatkozik lényegében minden terjesztésre szánt és jel-
lemzően kéziratos formában ránk hagyományozott szövegre: a
költészeten és a mai értelemben vett „szépirodalmon” túl például
történetírói munkákra, szónoki beszédekre, filozófiai és tudomá-

⁷ A továbbfejlesztett *Index Thomisticus* ma már szabadon elérhető az interneten: www.corpusthomicum.org.

⁸ Bővebb áttekintés olvasható a TLG honlapján: stephanus.tlg.uci.edu/history.php, lásd továbbá Brunner 1993 és Crane 2004 már említett tanulmányait.

nyos szövegekre is. Ami tehát kimaradt a TLG-ből, az a késő antik görög irodalom, a nem irodalmi papiruszok (szerződések, magánlevelek stb.), illetve a feliratok.⁹

A nagy mennyiségű szöveg rögzítésének megkezdése előtt a rendszer infrastruktúráját is meg kellett tervezni és teremteni. Szükség volt az ógörög szövegekben, illetve azok kritikai kiadásiban nagy számban előforduló diakritikus és egyéb textuális jeleket is rögzítő kódolási rendszerre (ez a Beta Code nevet kapta),¹⁰ valamint az adatok tárolását és keresését lehetővé tevő speciális szoftverre és hardverre (ezek együttesét az archaikus görög költők egyike után Ibycusnak nevezték el). A digitális bölcsészet jelene szempontjából sem mellékes, hogy a bölcsészek már akkor sem pusztán megrendelőként fordultak az informatikusok felé a digitális eszközök kifejlesztésekor: a két csoport között személyi átfedések is voltak. Az Ibycus rendszer kifejlesztője, az első digitális konkordanciák kapcsán már említett David Packard mérnök és ókortudós volt egy személyben, s nem melleleg a számítástechnika történetében fontos szerepet játszó Hewlett–Packard cég társalapítójának fia; a TLG általános szövegkorpusszá fejlesztésének ötlete (s a szükséges anyagi források egy része) pedig az elektronikai eszközöket fejlesztő Zenith Corporation alapítójának lányától, Marianne McDonaldtól származott, aki éppen a boldogság kifejezéseit vizsgálta Euripidészről szóló doktori disszertációjában. David Packard nem sokkal később, az 1980-as években tervezett egy speciális személyi számítógépet is a TLG használatához, ez azonban az általános célú PC-k terjedésével hamar idejé múlttá vált. Maga az adatbázis idővel mágnesszalagról CD-ROM-ra költözött; 1985-ben éppen a TLG lett az első publikált, nem zenei adatot tartalmazó CD. A David Packard alapította Packard Humanities Institute eközben elkészítette a

⁹ A görög nyelvű feliratok viszonylag bő, de korántsem teljes adatbázisa: epigraphy.packhum.org; a papiruszgyűjtemények tekintetében lásd papyri.info.

¹⁰ A Beta Code rövid áttekintése és teljeskörű dokumentációja egyaránt elérhető: stephanus.tlg.uci.edu/encoding.php.

TLG latin megfelelőjét, a PHI adatbázist, mely szintén Kr. u. 200-ig tartalmazta a római irodalmi szövegek lényegében teljes korpuszát. E két, technikailag is azonos felépítésű és ugyanazon kliensprogramokkal kezelhető adatbázis tehát a 90-es években egy-egy CD-n a klasszikus görög és latin irodalom túlnyomó részét hozzáférhetővé tette – még hozzá megbízható és sok esetben meglepően friss kritikai kiadások szövegét átvéve –, és az egész korpuszra kiterjedő, gyors keresési lehetőségeket biztosított.

Ezt követően viszont a két adatbázis útjai elváltak egymástól. A TLG 2001-ben webes felületet kapott, és előfizetéses üzleti modellre váltott (bár egy igen korlátozott szövegválogatás továbbra is szabadon elérhető), a kronológiai lefedettség pedig jelentősen bővült: ma már a bizánci irodalom is elérhető egészen 1453-ig, így a TLG összesen mintegy százmillió szó terjedelmű szöveget tartalmaz.¹¹ Integráltak több, a szövegek értelmezését segítő eszközt: a sztenderd ógörög–angol nagyszótár legfrissebb kiadását, valamint egyes speciális szerzői szótárakat. Megtörtént a szövegek lemmatizálása (a ragozott szóalakok megfelelő lexémához rendelése), így könnyen megtalálható egy-egy szó bármely előfordulása, függetlenül attól, hogy milyen alakban jelenik meg az egyes szövegekben – ez a görög nyelv rendkívüli morfológiai és dialektikus változatosságát tekintve különösen fontos. E fejlesztéseknek köszönhetően tehát egyetlen lekérdezéssel kideríthető, hogy egy adott jelentésmezőhöz tartozó görög szavaknak, vagy egy adott jelzős kifejezésnek bármely ragozott alakja (szórendtől függetlenül) hol fordul elő a teljes korpuszban, vagy azon belül egy szerzői életműben, irodalomtörténeti korszakban, esetleg műfajban. Újabban a TLG ún. n-gram keresője arra is képessé vált, hogy két kiválasztott szöveget összehasonlítva automatikusan felismerjen egyes nem szó szerinti idézeteket, illetve mindkét szövegben előforduló idiómákat is.

Miközben a TLG így fejlődött, a latin PHI adatbázis hosszú időre tetszhalott állapotba került, míg végül 2015-ben készítői

¹¹ stephanus.tlg.uci.edu.

ingyenes online felületen tették elérhetővé a legutolsó, 1998-ban megjelent verziót.¹² Az adatbázis tehát nem bővült, s azóta sem bővül új szövegekkel,¹³ integrált szótárakkal, s a felhasználói felület is – noha sok szempontból praktikus – igen korlátozott keresési lehetőségeket nyújt. A TLG-nek ma a korpusz kiterjedtségét és a keresési lehetőségeket tekintve inkább a Brepols kiadó 2009 óta elérhető *Library of Latin Texts* (LLT) nevű, ugyancsak előfizetéses adatbázisa feleltethető meg, mely az antik latin nyelvű szövegek lényegében teljes (tehát nem csak a Kr. u. 200 előtti) fennmaradt korpuszán túl hozzáférést biztosít rengeteg középkori, újkori és kortárs szöveghez egészen a II. vatikáni zsinatig bezárólag, s jól paraméterezhető, lemma-alapú keresővel rendelkezik.¹⁴

3. A digitális korpuszok és a nagyközönség

A fentiekben tárgyalt három szövegkorpusz kifejlesztéséhez a klasszika-filológia szaktudományos igényei vezettek, ez pedig az adatbázisok és keresőfelületük jellegét is alapvetően befolyásolja. Célközönségüket maguk a klasszika-filológusok alkotják: azok a latin és ógörög nyelvben, valamint az antik kultúrában és történelemben jártas szakemberek, akiknek a szövegek szó szoros értelmében vett, elsődleges megértéséhez jellemzően nincs szükségük fordításokra, az egyes szóalakok nyelvtani magyarázatára, és szótárhoz sem minden egyes mondat értelmezésekor

¹² latin.packhum.org (lásd továbbá az online változatról írt recenziókat: Loar 2017; Kozák 2018).

¹³ A PHI-t (is) kiegészítő speciális korpusznak tekinthető a késő antik, elsősorban nem keresztény szövegeket egybegyűjtő digilibLT: digilibLT.lett.unipmn.it. A latin nyelvű feliratok legteljesebb, teljes szövegű keresést biztosító online metakeresője az Epigraphik-Datenbank Claus/Slaby (db.edcs.eu; Adamik Béla szíves szóbeli közlése).

¹⁴ clt.brepolis.net/llta. Az LLT egy keresztény latin szövegeket tartalmazó CD-ROM adatbázis (CETEDOC) alapjaira épült.

kell nyúlniuk. Érdeemes röviden két olyan szöveges adatbázisról is szót ejteni, mely (részben legalábbis) más közönséget céloz meg. Az egyik a Latin Library, mely amatőr vállalkozásként már a kilencvenes évek végén – amikor még a PHI adatbázis is csak CD-ROM-on volt elérhető – szabadon hozzáférhetővé tett az interneten rengeteg antik és posztantik latin szöveget.¹⁵ A mai szemmel nézve rendkívül egyszerű honlapnak nincs saját keresője; a szerzői jogi korlátok miatt régi és sok esetben eleve megbízhatatlan szövegkiadásokon alapul, ráadásul a digitalizálás során keletkezett szöveghibák egy része sincs javítva. A Latin Libraryt tehát nem tekinthetjük megbízható tudományos eszköznek (erre a latin szövegeket maguk számára gyakran innen nyomtató egyetemi hallgatókat is figyelmeztetnünk kell). A másik adatbázis az 1995 óta elérhető és folyamatosan fejlesztett Perseus Digital Library,¹⁶ mely ógörög és latin szövegeket egyaránt tartalmaz, vagyis ahelyett, hogy az antik irodalom két nagy korpuszát egymástól nyelvi alapon elválasztaná, sokkal inkább azt hangsúlyozza, hogy azok sok tekintetben kulturális egységet alkotnak, ezért vizsgálhatuk elválaszthatatlan egymástól. Legalább ennyire fontos, hogy a Perseus hozzáférhetővé tesz fordításokat, szótárakat, kommentárokat, valamint egy-egy szaklexikont is, melyek vonatkozó szakaszai a tárgyalt antik szövegek mellett is megjeleníthetők, az antik szövegek megértését pedig az egyes szavak nyelvtani analízise is segíti. A Perseus tehát nem csupán szövegtörzs, hanem „digitális filológiai munkakörnyezet” (vagy annak prototípusa) is egyben, mely sok esetben a tudományos munkát is megbízhatóan támogatja. Kifejezetten szövegtörzsként viszont hagy kívánnivalót maga után: a már tárgyalt korpuszokkal (TLG, PHI, LLT) ellentétben az ógörög és római (vagy latin) irodalom egyetlen korszakára vonatkozólag sem tekinthető teljesnek, vagyis a segít-

¹⁵ www.thelatinlibrary.com; lásd még Gellar-Goad 2016 recenzióját, továbbá Tarrant 2016: 149–151 kritikus megjegyzéseit.

¹⁶ www.perseus.tufts.edu; recenzió erről: Lang 2018. A Perseus legújabb, 5.0-ás verziója megjelenés előtt áll; fenti megjegyzéseim és Lang recenziója még a 4.0-ás verzió alapjának.

ségével végrehajtott keresések sem lehetnek minden tekintetben reprezentatívak.

E helyütt mégsem a Latin Library és a Perseus hiányosságait szeretném hangsúlyozni, hanem azt, hogy a digitális szövegtörzsek műfaján belül mindkettőnek meghatározó szerepe van a klasszika-filológia társadalmi beágyazottságának, a szakmán kívüliekkel való kapcsolatának erősítésében. Az antik irodalmi szövegeket nem elzárt, a kutatók számára hozzáférhető és vizsgálható törzsként mutatják be, hanem mindannyiunk közös és éppen ezért szabadon hozzáférhetővé teendő kulturális örökségként. Ennek az örökségnek a megértése nehéz ugyan és klasszika-filológus szakértőket igényel, de utóbbiak feladata az is, hogy a megértés legalább egyes útjainak bejárásában az érdeklődők lehető leg szélesebb köre számára segítséget nyújtsanak.

4. A törzsek korai kifejlesztésének néhány mellékhatása

A klasszika-filológusok joggal büszkélkedhetnek azzal, hogy rendkívül korán hozzákezdtek a digitális szövegtörzsek kifejlesztéséhez; ritkábban esik azonban szó ennek néhány káros mellékhatásáról. E mellékhatások részben technikai jellegűek. Feltehetőleg már az adatbázisok (nem minden esetben jól dokumentált) felépítése sem ideális a mai elvárások felől nézve – különösen igaz lehet ez az 1998 óta lényegében nem fejlesztett PHI-re. Még feltűnőbb azonban az elmaradás az egyes szövegek digitális reprezentációját tekintve. A TLG és a PHI tervezésekor még nem léteztek a szövegek digitális, filológiai igényű kódolásának olyan, ma lényegében megkerülhetetlen szabványai, mint a Unicode, a TEI és az XML. A Beta Code (lásd fent) karakterkódolását viszonylag egyszerű Unicode-dá konvertálni, de más elemeinek TEI/XML-re való átalakítása nem automatizálható, és sok esetben a Beta Code nem is rögzít minden olyan információt, mely nélkül a TEI/XML kódolás

nem valósítható meg.¹⁷ A meglévő korpuszok technikai szempontból vett korszerűsítése tehát hatalmas emberi és anyagi erőforrásokat igényelne.

Kérdés ugyanakkor – amint azt Franz Fischer nemrég megjelent tanulmányában vizsgálja –, hogy reális-e ugyanolyan elvárásokat támasztani a digitális szövegtörzsekkel, mint egyes szövegek digitális (értsd: eleve digitális közegben született, *born digital*) kritikai kiadásával szemben.¹⁸ Értelemszerűen több időt és energiát szánhatunk egységnyi terjedelmű szöveg digitális kódolására, ha egyetlen mű kiadását készítjük el, mint ha több száz vagy ezer szövegből építünk törzset; az utóbbi esetben az egységes reprezentáció igénye felülírhatja az egyes szövegek jellegzetességeihez legjobban idomuló és legaprólékosabb kódolás iránti igényt. A Fischer által javasolt kompromisszum szerint a digitális szövegtörzsekben helyet kapó szövegek esetében a „digitális” kritériumok rugalmasan kezelendők, a filológiai kritériumokból viszont nem engedhetnek a szerkesztők.

A klasszika-filológia meglévő törzseinek minősége e kompromisszumos javaslat szerint sem kielégítő. Korai fejlesztésük káros mellékhatásai között ugyanis filológiai jellegűek is akadnak, s ezek közül minden bizonnyal éppen a kritikai apparátus hiánya a legsúlyosabb. A TLG és a PHI fejlesztésekor még nem zajlott le az az „új filológiához” köthető szemléletbeli változás, melynek eredményeképpen a szövegkiadók által elvetett szövegvariánsok is felértékelődtek, és az interpretáció szempontjából érdekesebbé váltak.¹⁹ Szerzői jogi és technikai problémák mellett

¹⁷ Kifejezetten technikai szempontból a Perseus fejlettebb, mint a TLG és a PHI (például az újabban feltöltött szövegek már TEI/XML szabvány szerint kódoltak), ám a törzs nagyságát tekintve nem veheti fel a versenyt az utóbbiakkal, s ezért nem is léphet a helyükre.

¹⁸ Fischer 2017: 275–287; a szerkesztői szerepről a digitális szövegtörzsek vonatkozásában lásd még Crane–Rydberg–Cox 2000.

¹⁹ Az „új filológiáról” magyarul lásd például a *Helikon* 2000/4-es tematikus számát; a szövegvariabilitásáról elméleti összefüggésben Déri–Keremen–Krupp–Tamás 2011 vonatkozó tanulmányait (219–432).

valószínűleg ebből is fakad, hogy a TLG és a PHI megbízható nyomtatott kritikai kiadásokon alapul ugyan, de kizárólag a szövegkiadó által megállapított főszöveget veszi át; a szövegkiadások digitalizálása a bevezetőkre és a kritikai apparátusra nem terjedt ki. A TLG történetében többször is előfordult már, hogy egy szövegnek új szövegkiadása került az adatbázisba, ezek azonban a régi helyére léptek, a két kiadás nem vált párhuzamosan megjeleníthetővé, illetve kereshetővé.²⁰ A korpuszok jelenlegi változatában tehát nemcsak az esetleges szövegvariánsokra nem lehet rákeresni, hanem minden szövegtörténeti információ, továbbá a hagyományozott szöveg inherens variabilitásának ténye is elfedésre kerül – ennek pedig különösen a klasszika-filológusok legújabb, immár digitális szövegeken és szövegkorpuszokon nevelkedő, s azokat nem annyira a nyomtatott kiadások mellett, hanem inkább helyettük használó generációjára gondolva lehetnek módszertani kockázatai.²¹

²⁰ Vö. Crane–Bamman–Jones 2013: 35.

²¹ A Lipcsei Egyetemen működő *Open Greek and Latin* projekt (www.dh.uni-leipzig.de/wo/projects/open-greek-and-latin-project) egyik célja éppen az, hogy egy-egy szövegnek több szövegkiadását is digitalizálja, és azok főszövegét összehasonlítva legalább részleges rálátást biztosítson a szövegtörténetre is. Az antik és posztantik latin nyelvű költészet speciális korpusza, a *Musisque Deoque* (www.mqdq.it) célja, hogy a nyomtatott kiadások kritikai apparátusát is legalább részben digitalizálja: azokra az olvasatokra kiterjedően, melyek ugyancsak értelmes, az irodalmi interpretáció szempontjából figyelembe vehető szöveget eredményeznek. A szövegtörténeti információk jelenleg csak egyes szövegek (például Vergilius *Aeneide*) esetében érhetők el. E helyütt említendő továbbá, hogy az egyesült államokbeli Society for Classical Studies új digitális kritikai kiadássorozatot tervez indítani (digitallatin.org), melynek technikai és filológiai alapelvei jelenleg még kidolgozás alatt állnak (digitallatin.github.io/guidelines/LDLT-Guidelines.html).

5. A klasszika-filológia „hagyományos” gyakorlata és a digitális szövegtörzsek

Klasszika-filológusként csábító lehetőség természetesen, hogy azt mondjam: kollégáim valamiféle jóstehetségtől indítva kezdtek hozzá már a hetvenes években az említett digitális törzsek létrehozásához. Az okok azonban minden bizonnyal ennél praktikusabbak voltak, illetve részben a klasszika-filológia hagyományos módszertanából fakadtak. Lássuk először a praktikus okokat. Az antik (és azon belül is különösen a Kr. u. 200-nál korábbi, a TLG- és a PHI-törzsekben eredetileg összegyűjtött) latin és ógörög irodalmi szövegek közepes méretű és statikus törzset alkotnak. Közepes méretben e helyütt azt értem, hogy a törzs egyrészt van annyira terjedelmes, hogy a filológus egyéni, természetes memóriája már ne tudja teljes egészében befogni, sőt, a nyomtatott szótárak, konkordanciák és egyéb kézikönyvek is csak korlátozottan legyenek képesek reprezentálni, egy-egy jelenség összes előfordulása helyett csupán példákat hozva. A törzs ugyanakkor – egy-egy élő nyelvi irodalommal összehasonlítva – eléggé korlátozott terjedelmű is ahhoz, hogy a szövegek kézi felvitele (az ellenőrzést is beleértve) ugyan hatalmas, de mégsem beláthatatlan mennyiségű munkát jelentsen. A szóban forgó törzs továbbá majdnem teljesen statikus. Új antik szövegek értelem szerűen nem keletkeznek, s bár időről időre még mindig előkerülnek eddig ismeretlen irodalmi szövegek (gondoljunk például a kétezres évek közepén publikált Szapphó-papirusztörzsekre), ezek a törzs összterjedelmét érdemben nem módosítják. A törzs statikus jellegéből fakadóan nem kellett és ma sem kell attól tartani, hogy a digitalizálás nem tud majd lépést tartani a törzs bővülésével. Az említett adatbázisok kifejlesztése tehát meglévő tudományos igények kielégítését ígérte, ugyanakkor reális célnak is bizonyult. Összehasonlításképpen: a hetvenes évek technikai infrastruktúrájával az angol vagy német nyelvű irodalmi szövegek teljes körű digitalizálása, akár csak a századfordulóig bezárólag, vagy akár csak egy-egy országra korlátozva minden bizonnyal elképzelhetetlen lett volna.

A digitális szövegtörzsek korai kifejtésének másik oka, mint említettem, vélhetőleg a diszciplína hagyományos, digitális kort megelőző „analóg” hagyományában keresendő. A ma ismert antik görög és latin szövegek törzse nemcsak közepes méretű (a fenti értelemben), hanem jelentős mértékben fragmentált is. Nemcsak azokra a művekre gondolok, melyek töredékesen maradtak ránk, vagy akár kizárólag címüket ismerjük, hanem azokra is, melyek többé-kevésbé hiánytalanul átvészelték az évezredek – s ez utóbbiakra sem kizárólag azért utalok, mert maguk is többé-kevésbé bizonytalan kézirati hagyomány alapján rekonstruálhatók csupán. Mindazok az antik szövegek, fragmentumok és címek, melyeket ma ismerünk, csupán töredékek egy szöveges rommezőn: emlékeztetnek minket egy valaha sokkal terjedelmesebb törzsekre. Minden kultúrára igaz természetesen, hogy egyes szövegei elveszhetnek a jövő filológusai számára; de a helyreállíthatatlan szövegromlás és pusztulás mértéke az ókori kéziratok esetében értelemszerűen különösen nagy.

Ezt a fragmentált törzset ugyanakkor a klasszika-filológusok – bár töredékességével természetesen ők maguk vannak a leginkább tisztában – hagyományosan nagyon is egységes és egybefüggő szöveganyagként kénytelenek kezelni. A tőlünk nemcsak időben, hanem kulturális és nyelvi tekintetben is távoli szövegek értelmezéséhez – sőt gyakran elsődleges, grammatikai megértéséhez is – minden rendelkezésre álló összehasonlító adatra szükségünk van. Két, az irodalomtörténeti narratívákban egymástól igen távoli szöveget, illetve azok egy-egy részletét is közvetlenül összevetjük egymással (például egy Kr. e. 5. századi görög nyelvű orvosi szöveg néhány szavas töredékét Vergilius *Aeneis*-ének egy sorával), ha úgy véljük, hogy ez segít megoldani egy szövegkritikai problémát, megválaszolni valamilyen nyelvi vagy tartalmi kérdést, vagy bármilyen tekintetben előmozdítja az interpretációt. Ezzel természetesen nem feltétlenül állítjuk, hogy a két szöveg között irodalmi értelemben vett intertextuális kapcsolat is fennállna (azaz idézetről vagy allúzióról volna szó), s még kevésbé azt, hogy a későbbi szöveg szerzője ismerte a korábbi szöveget. Mégis, filológiai tevékenységünk a szó legáltalánosabb – ugyan-

akkor nagyon is pragmatikus – értelmében intertextuális, szövegközi kapcsolatokat hoz létre a korpusz különböző szövegei, pontosabban azok egy-egy jellemzően apró részlete között. Természetesen nem azt állítom, hogy kizárólag klasszika-filológiai gyakorlat volna ez, csupán azt, hogy az antik szöveg megértése és értelmezése esetében más filológiákkal összehasonlítva különösen jellemző praxisról beszélhetünk.

Nem véletlen, hogy az értelmezett szöveget mintegy atomjaira bontva – Roland Barthes szavaival: „csillagokká repesztve”²² – vizsgáló kommentár mindig is ennyire nélkülözhetetlen és jellemző szakirodalmi műfaj volt és maradt mindmáig a klasszika-filológiában.²³ Egy adott szövegrészhez – egyszerre gyakran csak egy-két szóhoz – írt tipikus kommentár-bejegyzés kisebb-nagyobb részben éppen az értelmezést segítő „párhuzamos szöveghelyek” (*loci paralleli/similes*) listájából épül fel. E listát gyakran a „vö. pl.” hírhedt formulája vezeti be, helyettesítve és a kommentár olvasójára bízva a párhuzamok interpretációját.²⁴ Ez a formula és a párhuzamos szöveghelyeket jelölő többé-kevésbé szabványos, minden kiadás által követett rövidítések (például Cic. *Leg.* 2.30 = Cicero, *De legibus* 2. könyv, 30. *caput* [fejezet]) mintegy pre-digitális linkek módjára²⁵ biztatják a kommentár olvasóját, hogy az egyik szövegről a másikra „ugorjon”, majd onnan vissza az elsőre, vagy éppen (a másik szöveghez írt kommentárbe-

²² Barthes 1997: 25–26.

²³ A kommentár műfaja kiemelt figyelmet kapott az utóbbi két évtized filológiatörténeti és -elméleti kutatásaiban. Lásd például a következő tanulmányköteteket: Most 1999, különös tekintettel Don Fowler tanulmányára; Gibson–Kraus 2002; Kraus–Stray 2016.

²⁴ Erről bővebben lásd Gibson 2002. Gibson a költői szövegek értelmezési hagyományára koncentrálna a párhuzamos szöveghelyek egy lehetséges tipológiáját is kidolgozza (333–346).

²⁵ A Classical Works Knowledge Base (cwkb.org) a *linked open data* technikájával lehetővé teszi e hagyományos szöveghely-jelölések digitális linkeké alakítását, a felhasználó pedig választhat, hogy melyik digitális korpuszban szeretné az adott szöveghelyet megnyitni. Lásd még a Canonical Text Services szolgáltatást (cite-architecture.github.io/cts/).

jegyzést is megnézve) tovább egy harmadikra, negyedikre és így tovább a végtelenségig. Az antik irodalom kutatása során tehát a klasszika-filológusok az ilyen értelemben vett szövegek közötti kapcsolatok végtelenül sűrű és komplex hálózatát is felépítették; és éppen ezek a kapcsolatok teszik az egyébként töredékes korpuszt – hangsúlyozottan mint a tudományos vizsgálódás tárgyát – a fent tárgyalt értelemben mégis egységessé és egybefüggővé. Ez a hálózat természetesen csak virtuális, hiszen egészében senki sem rajzolta még fel – a vállalkozás minden bizonnyal lehetetlen is volna –, de az antik szövegek tudományos recepciójáról minden bizonnyal nagyon sokat elárulna, ha látnánk, hogy mely szövegek (illetve szövegrészek) játszanak középponti szerepet benne: melyekre hivatkoznak különösen sokszor egy-egy másik szöveg magyarázataképpen, vagy fordítva, melyek esetében szokás átlagban felüli számú párhuzamot felvonultatni a magyarázathoz.²⁶

A digitális szövegek korpuszok korai kifejlesztése valószínűleg részben annak is köszönhető (ha ez nem is feltétlenül tudatosult, s jellemzően ma sem tárgyalt szempont), hogy azok jutnak a legközelebb az antik szövegek imént vázolt, hálózatként elképzelt modelljének gyakorlati megvalósításához. Ha a vizsgálatba bevonható szövegeket nemcsak közös fizikai térben (könyvtárban, könyvespolcon, de különálló kötetekben) helyezzük egymás mellé, hanem egyetlen, konzisztensen kódolt digitális szövegtérben (adatbázisban) helyezzük el őket, akkor új lehetőségek nyílnak meg a kutatók előtt. A legkézenfekvőbb haszon nyilvánvalóan a szövegekhez való hozzáférés megkönnyítése és felgyorsítása: a filológus a számítógépe előtt ülve néhány másodperc alatt az őt érdeklő szövegek bármelyikének bármely

²⁶ Az antik szövegeket a tudományos publikációkban összekapcsoló „hivatkozási hálózatok” automatikus digitalizálásának lehetőségeiről lásd Romanello 2016: 21–39. Vergilius *Aeneis*ére – és csupán a szakirodalom egy részletére – vonatkozólag lásd az „*Aeneid* in Jstor” (aeneid.citedloci.org/) szolgáltatást, mely az eposz „hőterképét” is felrajzolva mutatja, hogy egy-egy *Aeneis*-sorra mely Jstorban elérhető tanulmányok hivatkoznak (idézettel vagy a nélkül).

részletét meg tudja jeleníteni (és készülő tanulmányába átmásolni). A szövegről szövegre ugrás meggyorsításán túl azonban bizonyos típusú szövegekőzi párhuzamok keresésének is új útjai nyílnak meg. A legkönnyebben természetesen az azonos lexikai bázison alapuló szövegekőzi párhuzamok találhatók meg, illetve egy-egy szó vagy kifejezés nyelv- és irodalomtörténeti „életútja” vázolható fel;²⁷ de a lehetőségeknek csak az szab határt, hogy digitális szövegtörpuszunk keresője milyen paraméterek szerint „finomhangolható”, s hogy emberi közreműködéssel vagy automatizált, gépi eszközökkel mennyi és milyen (nyelvi, irodalmi, kulturális) információt kapcsolunk a szöveg egy-egy szegmenséhez (betűhöz, szóhoz, mondathoz stb.). A TLG kifejlesztésekor éppen a teljes korpuszra kiterjedő keresés, nem pedig a gyors hozzáférés biztosítása volt az elsődleges cél. Ez annak tükrében tűnik különösen említésre méltónak, hogy – egy nemrég megjelent tanulmány szerint – az angol filológia egyes hasonló digitális korpuszainak létrehozásakor a kilencvenes és kétezres években a szövegeken átívelő kereshetőség másodlagos szerepet játszott, s e funkció majdnem ki is maradt a fejlesztésből.²⁸

6. Az olvasás módozatai

A digitális szövegtörpuszok által nyújtott szolgáltatások nem egyszerűen felgyorsíthatják a korábban is végzett műveleteket, hanem – ezzel összefüggésben – át is alakíthatják a létező praxisokat, munkamódszereket. Egyszerű példával élve: egy engem érdeklő tanulmányban említett, de nem idézett szöveghelyek kö-

²⁷ Erre jó példa lehet John Richardson monográfiája az eredetileg „parancsnoki jogkör”, s csak később (a kora császárkortól) „birodalom” jelentéssel bíró *imperium* szó jelentésfejlődéséről írott, a korábbi munkáknál lényegesen több forráshelyet vizsgáló monográfiája (Richardson 2008); a szerző digitális munkamódszeréről bővebben lásd Richardson 2005: 139–140.

²⁸ Bilansky 2017: 513–515, az Early English Books Online és a Women Writers Online adatbázisokról.

zül minden bizonnyal többet fogok ténylegesen is megnézni, ha nem kell mindegyikért könyvtárba mennem, és ott adott esetben a raktárból kikérnem. Nem feltétlenül arról van tehát szó, hogy ugyanazt a munkát gyorsabban végzem el, hanem arról, hogy a munka elvégzésére szánt, szükségképpen véges időben reprezentatívabb áttekintést szerezhettek az adott téma szempontjából releváns forrásokról.

Az imént tudatosan választottam az „olvasni” helyett a „megnézni” és az „áttekintést szerezni” kifejezéseket. A digitális médiumok fejlődésével párhuzamosan növekvő figyelmet kaptak a szöveges tartalmak befogadásának, tág értelemben vett olvasásának különböző módozatai és stratégiái. Az (irodalmi) szöveg aprólékos, a lehető legtöbb részletre kiterjedő és ezeket a részleteket az esztétikailag egységesnek tekintett mű egészének interpretációjában hasznosító lineáris „szoros olvasás” (*close reading*) hosszú időn át megkérdőjelezhetetlen stratégiája mellett – vagy éppen helyett – ma olyan módszerek és jellemzően digitális eszközeik képezik kiemelt figyelem tárgyát, mint a makroanalízis, a „távoli olvasás” (*distant reading*) és az „algoritmikus olvasás”.²⁹ A különböző kutatók által javasolt – részben egymást kiegészítő, részben egymással rivalizáló – terminusok, illetve az általuk jelölt értelmezői stratégiák közös jellemzője, hogy az „olvasás” tárgyaként az emberi elme által terjedelmük miatt nehezen vagy egyáltalán nem belátható szövegeket, illetve szövegkorpuszokat határozzák meg, az „olvasás” digitális eszközeire koncentrálnak, és sok esetben a szövegek „szabad szemmel”, a szoros olvasás stratégiáját követve nem is tudatosuló, inkább grafikonokon és térképeken ábrázolható tulajdonságainak jellemzően kvantitatív elemzésére

²⁹ Makroanalízis: Jockers 2013; távoli olvasás: Moretti 2005 és 2013; algoritmikus olvasás: Ramsay 2011. Az olvasási praxisok átalakulásáról általában lásd Hayles 2012: 55–79; a távoli olvasásban rejlő lehetőségekről magyarul lásd például Labádi 2014; Péter 2016; vö. továbbá a *Publications of the Modern Language Association of America* 2017 májusában megjelent 132 (3)-as számának tematikus – és Moretti megközelítésével szemben több esetben kritikus tanulmányokat tartalmazó – összeállítását.

építik az interpretációt. Franco Moretti provokatív hipotézise szerint az irodalmi kánont alkotó viszonylag kevés mű szoros olvasása helyett a teljes irodalmi produkció „távoli olvasása” révén tudunk egy-egy műfajról vagy irodalomtörténeti korszakról releváns értelmezéseket adni – Moretti például az angol regény és alműfajai 18–19. századi történetét elemzi több ezer regény címét és narratív alapsémáját (de hangsúlyozottan *nem* a szó szoros értelmében vett szövegét) vizsgálva.³⁰

Az ilyen értelemben vett „távoli olvasás” lehetőségei a klasszika-filológiában korlátozottnak tűnnek, nem utolsósorban éppen az antik irodalmi szövegek korpuszának fentiekben már tárgyalt közepes mérete és fragmentáltsága következtében. A könyvnyomtatás és az irodalmi tömegtermelés kezdete előtt keletkezett szövegekről van szó, melyek közül sok mára végérvényesen elveszett. A szövegek összességüként definiált korpusz még eléggé terjedelmes ahhoz, hogy kvantitatív, statisztikai megközelítésben vizsgáljuk például egy-egy szó, kifejezés vagy akár grammatikai szerkezet előfordulásait; ám ha például az egyes műfajokba sorolható és legalább címként fennmaradt művek számát vesszük alapul, akkor már nem kapunk megfelelően nagy elemszámú adathalmazt ahhoz, hogy olyan „távoli olvasásnak” vessük alá azt, mint Moretti tette az angol regénnyel. A klasszika-filológia szempontjából ezért azok a digitális olvasás módozataival összefüggő kutatások tűnnek különösen relevánsnak, melyek a *full text* adatbázisokat mint szöveges „műfajt”, s ezzel összefüggésben az adatbázis „olvasását”, kezelését mint filológiai praxist vizsgálják. E kutatások közül több is hangsúlyozza, hogy az egyedi szöveg szoros olvasása és az adatbázison lefuttatott keresések eredménylistáinak a távoli olvasás gyakorlatával rokonítható áttekintése és elemzése inkább egymást kiegészítő, mintsem szembeállítható és önállóan alkalmazandó stratégiák, s hogy éppen az adatbázis mint műfaj az, ami elősegíti a két értelmezői stratégia kooperatív alkalmazását.³¹ A digitális szövegtörzsek kiváló eszközök arra,

³⁰ Moretti 2005: 3–33.

³¹ Hayles 2012: 175–247; Bilansky 2017: 518–519.

hogy egyszerre kapjunk nem negatív értelemben véve felületes áttekintést³² egy-egy nyelvi jelenség vagy téma előfordulásairól az adott korpuszban (s ezt mint kvantitatív információt akár listákban, grafikonokon stb. ábrázoljuk is), valamint e tevékenység közben egyúttal ki is válasszuk a találatok közül azokat, melyeket a szoros olvasás módszerével részletesebb (kvalitatív) vizsgálatra is érdemesnek tartunk. Hogy a *distant reading* és a *close reading* egymást kiegészítve alkalmazható értelmezői stratégiák, arra jó példát szolgáltathat a klasszika-filológia fent vázolt jellemző – de természetesen nem kizárólagos – értelmezési gyakorlata is. A szövegek közötti hasonlóságok kimutatása és ezáltal a szövegek komplex hálózatának építése során az adott esetben digitális szövegkorpuszok és azok keresői segítségével megtalált szöveges párhuzamokat mint „adatokat” jellemzően egy adott szöveg értelmezésében, annak szoros olvasásához kapcsolódóan használjuk fel.

7. Szövegpárhuzamok keresése digitális eszközökkel

Amint arról fent már esett szó, a szövegpárhuzamok keresése a digitális kort megelőzően is a klasszika-filológia jellemző gyakorlata volt. Kérdés tehát, hogy milyen tekintetben nyújtanak új lehetőséget erre a digitális szövegkorpuszok és keresőik. Először is természetesen abban, hogy több párhuzamot találhatunk a segítségükkel; csak hogy a párhuzamos helyek sokszor már így is kényelmetlenül hosszú listájának³³ további bővítése egy ponton túl

³² A „hiperolvasásnak” többek a digitális keresők által összeállított találati listák feldolgozásában hasznos technikáit tárgyaló tanulmányában James Sosnoski nyolcat különböztet meg; ezek közül a digitális filológia szempontjából a szűrés (*filtering*), a felületes (*skimming*) és válogató olvasás (*pecking*), a szövegek közti határok átlépése (*trespassing*) és a fragmentálás (*fragmenting*) tűnnek különösen fontosnak (Sosnoski 1999: 163–172).

³³ A „párhuzamvadászat” negatív aspektusairól lásd Gibson 2002: 347–353.

minden bizonnyal nem járna együtt a filológiai megértés elmélyítésével, s a digitális kutatás során ellen is kell állnunk a pusztá halmozás kísértésének. A hangsúlyt inkább a változatosságra helyezném. Olyan párhuzamokat is könnyedén megtalálhatunk a digitális keresők – és persze helyesen megválasztott keresőkifejezések – segítségével, melyekre korábban saját olvasottságunk, emlékezetünk szükségszerű korlátai, a szótárak és más kézikönyvek korlátozott példaanyaga miatt nem bukkantunk volna rá, vagy éppenséggel azért nem, mert akár tudatosított, akár tudat alatti előzetes értelmezési gesztussal bizonyos szövegeket eleve nem találtunk relevánsnak a komparatív vizsgálat szempontjából. Ha Vergilius *Aeneis*ének egy részletét vizsgálva a tíz, már korábban is felismert homéroszi párhuzamot további öttel egészítjük ki, azzal viszonylag keveset nyertünk; de ha találunk akár csak egy korábban nem dokumentált releváns párhuzamot például a jogi szövegek között, az lényegesen módosíthatja a vizsgált szövegrészről alkotható értelmezéseket. A digitális szövegkorpuszokon végrehajtott keresések többek között éppen arra bizonyulhatnak különösen alkalmasnak, hogy a korábbi kutatásban rögzült, esetenként egyoldalú értelmezéseket újraérteljük és kiegészítsük egy-egy szöveghely intertextuális háttérének teljesebb feltárással.³⁴

A változatosság mellett a reprezentativitás is a digitális korpuszok használata mellett szólhat. A nyomtatott szótárak (a rendkívül ritka szavaktól eltekintve) nem sorolhatják fel egy-egy szó vagy akár szókapcsolat összes előfordulását. A filológusok gyakran relatív és pontatlan kijelentéseket kénytelenek tenni arról egy-egy szöveghely kapcsán, hogy az adott kifejezés „gyakran/ritkán fordul elő a klasszikus latinban” vagy „a költői szövegekben”. De vajon mi volna az átlagos: nem is túl gyakori és nem is túl ritka előfordulás? És tételesen mely szövegek képezik a „klasszikus

³⁴ Ennek lehetőségeit egy készülő esettanulmányban mutatom be részletesen, egy kulcsfontosságú *Aeneis*-helyhez (*imperium sine fine dedi*, „háttartalan/végtelen hatalmat adtam [a rómaiaknak], *Aen.* 1.279) kapcsolódóan: Kozák (megjelenés előtt).

latin” vagy a „költői szövegek” korpuszát? A digitális eszközök lehetővé teszik, hogy pontosítsuk ezeket a kijelentéseket: „az adott kifejezés X alkalommal található meg az Y digitális korpuszban” vagy annak „Z alkorpuszában”.³⁵ A hasonló kijelentések nem is annyira önmagukban, mint inkább egymással összehasonlítva sugallhatnak a korábbiaknál megalapozottabb értelmezéseket egyes kifejezések relatív gyakoriságáról.

A fentiekkel összefüggésben fontosnak tűnik a digitális korpuszokban lefuttatott keresések reprodukálhatósága, illetve ennek a publikálás során való biztosítása is. A filológusnak legalább a kulcsfontosságúnak ítélt esetekben érdemes pontosan dokumentálnia, hogy milyen keresőkifejezést és milyen keresési paramétereket használt. Az ideális eset az volna, ha a digitális szövegkorpuszok lehetőséget adnának arra, hogy a keresésekhez egyedi azonosítókat rendeljünk, s ezeket mint hivatkozásokat tüntessük fel publikációinkban, így megadva a lehetőséget az olvasók számára, hogy megvizsgálhassák ugyanazokat az adatsorokat, és esetleg más következtetésekre jussanak velük kapcsolatban.³⁶

Nem állíthatjuk ugyanakkor, hogy a digitális korpuszok segítségével talált párhuzamok – még ha számosabbak és reprezentatívabbak is – szükségképpen objektívebb képet is rajzolnának egy-egy adott szövegrész tág értelemben vett intertextuális kontextusáról. Fontos tudatosítani, hogy digitális keresésünk első lépése, a keresés tárgyát képező szó vagy kifejezés kiválasztása és ennek tükrében a keresőkifejezés összeállítása már eleve interpretatív aktus, és interpretatív jellegű a keresés eredményeképpen kapott találati lista elemzése is. A teljes találati lista csupán nyers adatokat tartalmaz, melyeket egyenként kell ellenőriznünk, értékelnünk és értelmeznünk. A lista tartalmazhat egyértelműen

³⁵ Vö. Crane–Bamman–Jones 2013: 53.

³⁶ E tekintetben a PHI-korpusz felhasználói felületének minimalizmusa kifejezetten hasznosnak mondható: minden egyes keresés, sőt minden megjeleníthető tartalmi egység saját URL-lel rendelkezik. Az LLT-adatbázisból egyedi formátumú fájlként letölthető egy-egy keresés összes paramétere.

téves találatokat: ilyenek lehetnek a homonímák, vagy a roszszul (esetleg a keresőrendszer korlátai miatt kényszerűen) választott keresőkifejezés következményei. A latin és az ógörög nyelv morfológiai jellemzői miatt ezeket a téves találatokat a digitális korpuszok jelenlegi verziói nem tudják automatikusan kiszűrni. A lista részét képezhetik továbbá a szó szoros értelmében véve nem téves, a kérdésfeltevésünk szempontjából mégis irrelevánsnak ítélt találatok is. Kereshetjük például két szó együttes előfordulásait, csak hogy a latin és az ógörög viszonylag szabad szórendje miatt ez nem feltétlenül jelenti azt, hogy a keresett szavak közvetlenül egymás mellett fordulnak elő egy mondatban. Ki tudom gyűjteni azokat a mondatokat, melyekben mindkét szó szerepel, ám a találatokat egyenként megvizsgálva tudom csak eldönteni a kontextus, a két keresett szó egymástól való távolsága és szintaktikai viszonya tükrében, hogy csupán esetleges vagy pedig jelentésképző erővel bíró jelenségről van-e szó, vagyis az értelmezés szempontjából is beszélhetünk-e „együttes előfordulásukról”.

A téves és az irrelevánsnak ítélt találatok kiszűrése után a lista gyakran még mindig viszonylag hosszú és heterogén marad: minél gyakoribb egyik vagy mindkét szó előfordulása a teljes korpuszban, annál inkább. Ilyen esetekben a találatok kategorizálása válik kulcsfontosságú értelmezői műveletté. Listánk tartalmazhat egyrészt olyan párhuzamokat is, melyeket a két szöveg között közvetlen kapcsolatot létrehozó idézetként vagy allúzióként azonosítunk: ezek az irodalmi intertextualitás minősített esetei. Egy-egy keresés lefuttatásakor a filológus valószínűleg abban reménykedik, hogy ilyeneket is felfedez majd – de persze éppen ezek esetében van viszonylag nagyobb esély arra is, hogy a korábbi szakirodalomban már dokumentált és interpretált párhuzamokról van szó, különösen, ha irodalomtörténeti szempontból kézenfekvő a két szöveg közti kapcsolat vizsgálata. E helyütt nem is a találatok ezen kategóriáját emelném ki, hanem azt a „szürke zónát”, amely az egyértelműen téves/irreleváns találatok és az egyértelműen releváns, az irodalmi interpretációt döntő mértékben alakító idézetek/allúziók között helyezkedik el. Eb-

ben a szürke zónában a „valamilyen szempontból potenciálisan releváns” találatok helyezkednek el. A digitális eszközökkel végzett kutatás módszertanát illetően sok szempontból ezek lehetnek a legérdekesebbek, éspedig éppen azért, mert nehéz a skála végpontjait képező másik két kategória egyikében elhelyezni őket. Segítségükkel jól tesztelhetjük – és kényszerítenek is arra, hogy teszteljük – az intertextualitás természetéről alkotott implicit vagy explicit elképzeléseinket, elméleteinket.³⁷ Hol vannak az értelmezhető intertextualitás határai? Az irodalmi allúzióktól az irreleváns találatok felé elindulva vajon hol lesz az a határ az említett szürke zónában, ahol még éppen megfogalmazhatónak, illetve már éppen megfogalmazhatatlannak tartjuk, hogy a kereső által adott találatok külön-külön vagy együtt milyen módon befolyásolhatják a digitális eszközökkel vizsgált szöveg értelmezését irodalmi, nyelvi, kulturális szempontból?

A digitális korpuszoknak és keresőknek az a fentiekben még másodlagosnak tekintett sajátossága, hogy a nem digitális módszerekhez képest több párhuzamos helyre hívhatják fel a kutató figyelmét, ebben az összefüggésben bizonyulhat mégis kifejezetten fontosnak. Vannak ugyanis olyan szövegpárhuzamok, melyek egyenként vizsgálva nem tűnnek különösebben relevánsnak – nem volna meggyőző állítás tehát, hogy a két szóban forgó szöveg között közvetlen és lényegi kapcsolat van –, ha viszont felismerjük, hogy a párhuzamként kínálkozó kifejezés ismétlődően fordul elő a korpuszban, akkor ezek a szöveghelyek együttesen mégis elárulhatják az elsődlegesen vizsgált szövegről, illetve annak kérdéses részletéről azt, hogy milyen műfaji, nyelv-

³⁷ Ebből a szempontból különösen feltűnő és sajnálatos, hogy az ezredforduló környékén (tehát már a digitális szövegtörzsek létrejöttét követően, de még jellemzően azok online elérhetővé válását megelőzően) megjelent, az antik irodalmi szövegek intertextualitását, illetve az intertextuális értelmezés klasszika-filológiai gyakorlatait általában vizsgáló monográfiák vagy egyáltalán nem tárgyalják a digitális szövegtörzseket (mint Hinds 1998), vagy csupán egy-egy mondatban tesznek említést róluk (így például Edmunds 2001: 22).

vi, kulturális regisztert szólaltat meg, milyen diskurzus(ok)hoz kapcsolható.

Az intertextualitás elméleteinek és gyakorlatának legradikálisabb, az egyes párhuzamok értelmezésétől legerőteljesebben elrugaszkodó teszteléséhez a digitális klasszika-filológiában jelenleg a Tesseract projekt szolgáltat eszközt.³⁸ Speciális, a plágiumkereső rendszerek elvén működő keresőjével nem egyes szavak, kifejezések előfordulásaira kereshetünk rá, hanem bizonyos paraméterek beállítása után a keresőrendszer két kiválasztott szövegből gyűjti ki a nyelvi hasonlóságot mutató részleteket, s ezeket bonyolult algoritmus szerint (elsősorban a hasonlóságot megalapozó szavak gyakorisága alapján) súlyozza is.³⁹ A találati listákon szereplő egyezések száma zavarba ejtő: Ovidius *Metamorphoses*ének 779 soros első éneke és Vergilius *Aeneis*e (összesen mintegy tízezer sor) között az alapbeállításokat változtatlanul hagyva nem kevesebb mint 6714 tételből álló találati listát kapunk, s e „párhuzamok” nagy részét minden bizonnyal nehéz lenne valamilyen irodalmi interpretáció során hasznosítani – gyakran az algoritmus által magas pontszámmal kiemelt egyezések is igencsak esetlegesnek tűnnek. A Tesseract célja végső soron nem is az, hogy egyes tételekkel gazdagítsa a dokumentált párhuzamok számát (bár sok idő és türelem birtokában éppenséggel erre is használható), hanem az, hogy a távoli olvasás eszközeivel mérhetővé és összehasonlíthatóvá tegye különböző szövegpárok „intertextuális távolságát”,⁴⁰ illetve ábrázolhatóvá tegye adott esetben e távolság relatív változásait – tehát a fenti példánál maradva például azt, hogy a *Metamorphoses* első énekében az *Aeneisszel* mint az epikus hagyomány meghatározó szövegével hasonlóságot mutató sorok

³⁸ tesseract.caset.buffalo.edu. A Tesseract korpusza nagyrészt a Perseus (lásd fent) latin szövegein alapul, ám annál lényegesen kisebb. A rendszer részletesebb bemutatását lásd: Coffee et al. 2013.

³⁹ A Tesseract súlyozási algoritmusáról lásd Forstall et al. 2015.

⁴⁰ Erről részletesen lásd Gawley–Diddams 2017.

gyakorisága többé-kevésbé állandó-e, vagy pedig valamilyen értelmezhető tendencia szerint változik.⁴¹

8. Összegzés

A tanulmányomban részletesen tárgyalt „hagyományos”, első generációs digitális szövegtörzsek és a Tesseræ intertextuális keresési lehetőségeit összehasonlítva kimondható, úgy vélem, hogy a klasszika-filológia immár mintegy fél évszázados „digitális történelme” jelenleg korszakhatárhoz ért. A TLG, a PHI és a hozzájuk hasonló korpuszok új eszközöket adtak a kutatók kezébe, s nem csupán könnyebbé, gyorsabbá tették munkájukat, hanem a kutatás módszertanára is hatást gyakoroltak – például azáltal, ahogyan a találati listák feldolgozásában a kvantitatív elemzési módszerek a korábbiaknál hangsúlyosabban egészítik ki a kvalitatív interpretációs stratégiákat. Az első generációs korpuszok és az általuk lehetővé tett keresések ugyanakkor, amint azt a fentiekben tárgyaltam, jól illeszkednek a klasszika-filológia hagyományos gyakorlatába: inkább továbbfejlesztik, illetve kiegészítik a korábbi praxisokat, mintsem radikális paradigmaváltást valósítanak meg. A digitális korpuszok megjelenése tehát a klasszika-filológia történetében döntő változás volt ugyan, mégis sok tekintetben evolúciós lépésként, az adott diszciplína saját történeti kontextusában a szerves fejlődés példájaként tekinthetünk rá.

Ezzel szemben a Tesseræ és a hozzá hasonló, fejlesztés alatt álló második generációs digitális eszközök az intertextuális kutatás terén inkább forradalmi változásokat ígérnek, illetve ilyenekkel fenyegetnek – nézőpont kérdése. Ma még egyáltalán nem látszik, hogy az általuk lehetővé tett és javasolt, a távoli olvasás

⁴¹ A projektben részt vevő filológusok legrészletesebb – a rendszer „kalibrálása” során készült – esettanulmánya (Coffee et al. 2012) a kora császárkori Lucanus *Pharsalia* című eposza első énekének, valamint Vergilius *Aeneis*ének intertextuális kapcsolatát vizsgálja.

stratégiáival rokonítható értelmezési módszerek hosszabb távon beilleszthetők lesznek-e a klasszika-filológia sztenderd, a kutatók által széles körben legitimként elismert praxisai közé. Abban azonban már most segíthetnek, hogy történeti kontextusba helyezzük a közelmúlt első generációs digitális eszközeit. Már látjuk nemcsak azok „analóg” előzményeit, hanem digitális utódjait is. A ma rendelkezésünkre álló informatikai infrastruktúra felől nézve a klasszika-filológia első generációs digitális korpuszai feltűnően egyszerű eszközök: egyrészt minél több (egy-egy korszakra kiterjedően lehetőleg az összes fennmaradt) latin és ógörög szöveget kell magukba foglalniuk, másrészt viszont informatikai szempontból jellemzően nagyon egyszerű, karakteralapú keresések végrehajtását teszik lehetővé az így létrejött korpuszon. Ha új eszközökként új praxisok kialakulásához vezettek is, a klasszika-filológusoknak a számítógép kezelésén túl nem kellett alapvetően új, szakmájuk hagyományaitól idegen készségeket elsajátítaniuk e korpuszok eredményes használatához. A Tesseræ keresési paramétereinek megfelelő beállításához, illetve az eredmények értelmezéséhez már legalább alapjaiban meg kell érteniük azokat az algoritmusokat, melyek szerint a kereső a találatokat súlyozza, a rendszer működését bemutató tanulmányok olvasója pedig matematikai képletekkel is találkozhat. A második generációs digitális eszközök tehát elődeiknél sokkal erőteljesebben kényszerítik ki az interdiszciplináris gondolkodást, más tudományok (ez esetben értelemszerűen a matematika és az informatika) szemléletmódjának alkalmazását. Akárhogyan alakuljon is hosszú távon a klasszika-filológia története a digitális korban, ez már önmagában is fontos hozzájárulás lehet ahhoz, hogy e diszciplína természetesen vehessen részt a digitális bölcsészet formálásában, illetve fenntartsa párbeszédképességét más, korunk tudományképét erőteljesebben meghatározó tudományokkal.

Irodalom

- Barthes, Roland 1997. *S/Z*. Budapest: Osiris Kiadó.
- Bilansky, Alan 2017. Search, Reading, and the Rise of Database. *Digital Scholarship in the Humanities* 32 (3): 511–527.
- Brunner, Theodore F. 1993. Classics and the Computer: The History of a Relationship. In Solomon, J. (ed.): *Accessing Antiquity: The Computerization of Classical Databases*. Tucson: University of Arizona Press. 10–33.
- Coffee, Neil – Koenig, Jean-Pierre – Poornima, Shakthi – Ossewaarde, Roelant – Forstall, Christopher – Jacobson, Sarah 2012. Intertextuality in the Digital Age. *Transactions of the American Philological Association* 142 (2): 383–422.
- Coffee, Neil – Koenig, Jean-Pierre – Poornima, Shakthi – Forstall, Christopher W. – Ossewaarde, Roelant – Jacobson, Sarah L. 2013. The Tesseract Project: Intertextual Analysis of Latin Poetry. *Literary and Linguistic Computing* 28 (2): 221–228.
- Crane, Gregory – Rydberg-Cox, Jeffrey A. 2000. New Technology and New Roles: The Need for “Corpus Editors”. *Proceedings of the Fifth ACM Conference on Digital Libraries*. New York: ACM Press. 252–253.
- Crane, Gregory 2004. Classics and the Computer: An End of the History. In Schreibman, Susan – Siemens, Ray – Unsworth, John (edd.): *A Companion to Digital Humanities*. Malden: Wiley–Blackwell. 46–55. [elérhető online: <http://www.digitalhumanities.org/companion>]
- Crane, Gregory – Bamman, David – Jones, Alison 2013. ePhilology: When the Books Talk to Their Readers. In Siemens, Ray – Schreibman, Susan (ed.): *A Companion to Digital Literary Studies*. Malden: Wiley, Blackwell. 29–64.
- Déri Balázs – Kelemen Pál – Krupp József – Tamás Ábel (szerk.) 2011. *Metafilológia 1: Szöveg – variáns – kommentár*. Budapest: Ráció.
- Edmunds, Lowell 2001. *Intertextuality and the Reading of Roman Poetry*. Baltimore, London: Johns Hopkins University Press.
- Fischer, Franz 2017. Digital Corpora and Scholarly Editions of Latin Texts: Features and Requirements of Textual Criticism. *Speculum* 92 (S1): 265–287.

- Forstall, Christopher – Coffee, Neil – Buck, Thomas – Roache, Katherine – Jacobson, Sarah 2015. Modeling the Scholars: Detecting Intertextuality Through Enhanced Word-level N-gram Matching. *Digital Scholarship in the Humanities* 30 (4): 503–515.
- Gawley, James O. – Diddams, A. Caitlin 2017. Comparing the Intertextuality of Multiple Authors Using Tesserae: A New Technique for Normalization. *Digital Scholarship in the Humanities* 32 (2): 53–59.
- Gellar-Goad, Ted H. M. 2016. Review: The Latin Library. *Society for Classical Studies Blog*: classicalstudies.org/scs-blog/ted-gellar-goad/review-latin-library.
- Gibson, Roy K. – Kraus, Christina S. (eds.) 2002. *The Classical Commentary: Histories, Practices, Theory*. Leiden, Boston: Brill.
- Gibson, Roy K. 2002. ‘Cf. e. g.’: A Typology of ‘Parallels’ and the Role of Commentaries on Latin Poetry. In Gibson – Kraus (eds.) 2002. 331–357.
- Hayles, N. Katherine 2012. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago: Chicago University Press.
- Hinds, Stephen 1998. *Allusion and Intertext. Dynamics of Appropriation in Roman Poetry*. Cambridge: Cambridge University Press.
- Jockers, Matthew L. 2013. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press.
- Kozák Dániel 2018. PHI Latin Texts (review). *RIDE. A Review Journal for Digital Editions and Resources* 8: ride.i-d-e.de/issues/issue-8/phi.
- Kozák Dániel (megjelenés előtt). The Intertextual Frontiers of Vergil’s „Empire Without Limit”: Digital Comments on *Aeneid* 1. 278–279.
- Kraus, Christina S. – Stray, Christopher (eds.) 2016. *Classical Commentaries: Explorations in a Scholarly Genre*. Oxford: Oxford University Press.
- Labádi Gergely 2014. Franco Moretti: Distant Reading (recenzió). *Irodalomtörténet* 95 (4): 561–564.
- Lang, Sarah 2018. Review of Perseus Digital Library. *RIDE. A Review Journal for Digital Editions and Resources* 8: ride.i-d-e.de/issues/issue-8/perseus.
- Loar, Matthew 2017. Review: The Packard Humanities Institute (PHI)—Classical Latin Texts. *Society for Classical Studies Blog*: clas-

- sicalstudies.org/scs-blog/matthew-loar/review-packard-humanities-institute-phi%E2%80%94classical-latin-texts.
- Moretti, Franco 2005. *Graphs, Maps, Trees: Abstract Models for a Literary History*. London, New York: Verso.
- Moretti, Franco 2013. *Distant Reading*. London, New York: Verso.
- Most, Glenn W. (ed.) 1999. *Commentaries – Kommentare*. Göttingen: Vandenhoeck & Ruprecht.
- Péter Róbert 2016. A Big Data kihívás és lehetőség a bölcsészettudományokban: digitális szövegek és metaadatok távoli olvasása. *Magyar Tudomány* 177 (11): 1323–1330.
- Ramsay, Stephen 2011. *Reading Machines: Toward an Algorithmic Criticism*. Urbana: University of Illinois Press.
- Richardson, John S. 2005. Indexing Roman Imperialism. *The Indexer* 24 (3): 138–140.
- Richardson, John S. 2008. *The Language of Empire: Rome and the Idea of Empire from the Third Century BC to the Second Century AD*. Cambridge: Cambridge University Press.
- Romanello, Matteo 2016. Exploring Citation Networks to Study Intertextuality in Classics. *Digital Humanities Quarterly* 10 (2). www.digitalhumanities.org/dhq/vol/10/2/000255/000255.html.
- Sosnoski, James J. 1999. Hyper-Readers and Their Reading-Engines. In Hawisher, Gail E. – Selfe, Cynthia L. (eds.): *Passions, Politics, and 21st Century Technologies*. Urbana: Utah State University Press. 161–177.
- Tarrant, Richard 2016. *Texts, Editors, and Readers: Methods and Problems in Latin Textual Criticism*. Cambridge: Cambridge University Press.

Nagy Károly Zsolt

Interaktív etnográfia Néprajzi kutatás interfészek közti terepen¹

1.

A különböző tudományos diszciplínáknak többnyire rendszeresen meg kell küzdeniük legitimitásukért, és ennek a küzdelemnek fontos ütközeteit vívják a korszerűségért. A korszerűségnek természetesen számos párhuzamos és egymást kiegészítő olvasata létezik. Az egyik oldalról nézve a korszerűség az éppen érvényes, a tudományág nemzetközi diskurzusait is meghatározó paradigmákhoz való illeszkedést, a „fősodorba” tartozó elméletek és módszerek reflexív alkalmazását jelenti. Egy másik oldalról nézve a korszerűség inkább egy olyan üzenet, melyet az adott diszciplína individuális vagy kollektív művelői igyekeznek közvetíteni saját magukról és munkájukról a nyilvánosság különböző szintjein. Egy magára valamit adó, magát anyagilag és erkölcsileg támogatásra érdemesként bemutatni akaró kutató vagy intézmény nem kerülheti el azt, hogy saját magát, munkáját és eredményeit saját társadalmi környezetének a korszerűségről alkotott elképzeléséhez szabva prezentálja, felruházza azokkal az attribútumokkal, melyekből a nagyközönség vagy a politikai és gazdasági döntéshozók az ő korszerűségére következtethetnek. Magyarán: nem elég korszerűnek lenni, annak is kell látszani.

Napjainkban a korszerűség egyértelmű attribútumai közé számítanak a mobilkommunikációs és számítástechnikai eszközök és az ezek segítségével működtetett digitális technológiák használata. Úgy jelennek meg a tudományos munka, az ahhoz

¹ A tanulmány létrejöttét a Bolyai János Kutatási Ösztöndíj támogatta.

kapcsolódó személyek és intézmények vizuális reprezentációján, ahogyan a 17. század németalföldi csendéletein a könyv: a korszerű tudás, a bölcsesség fogalmait vonják be a képet értelmező szemlélő gondolkodásába. A könyv szimbolikájához azonban egy idő után már nemcsak a bölcsesség, hanem a hiábavalóság, a „vanitas” is hozzátartozott. Ahogyan nőtt a nyomtatványok száma, és a bölcsesség mellett egyre nagyobb teret hódított magának a könyvek piacán a balgaság, úgy a könyv mint tárgy szimbolikus jelentése is egyre ellentmondásosabb lett.² Hasonló folyamat tanulni lehetünk a modern eszközök esetében is: ahogy terjednek, ahogyan az élet egyre több területét átformálják és egyre elérhetőbbé is válnak, s ahogyan ezzel párhuzamosan használatukban egyre többször jelenik meg a bölcsesség mellett a balgaság is, szimbolikájuk úgy válik egyre ellentmondásosabbá. Egyszer egy terepmunkám alkalmával, az egyik vidéki városunk központjában, egy nagy szupermarket előtt megfigyeltem egy férfit, akiről úgy tűnt, hogy mobiltelefonon tárgyal valakivel. Mindez abban az időben történt, amikor ez az eszköz még meglehetősen ritka és drága jószágnak számított, és birtoklását csak kevesek engedhették meg maguknak. A televízióban és a nyomtatott médiában viszont – nem is beszélve az óriásplakátokról – státusszimbólumként, illetve a modern életmód egyértelmű attribútumaként jelent meg. A meglehetősen elhanyagolt ruházatú férfi nagy hangon folytatta a beszélgetést, széles gesztusok kíséretében próbálta valamilyen jelentős összegű tartozás megfizetésére rábírní a vonal túlsó végén álló személyt. Ahogy közeledtem hozzá, egyre inkább a befolyásosság vagy legalábbis a helyi maffiavezérség képzetét társítottam az illetőhöz, míg mellé érve fel nem fedeztem, hogy az eszköz, amit a füléhez tartott, igazából nem mobiltelefon, hanem egy zsebszámológép volt. A számgombokkal és kijelzővel rendelkező, téglatest alakú tárgy a szóban forgó férfi számára – mint afféle „vesszőparipa”³ – kimerítette a „telefonság” minimális kritériumait, mert függetlenül attól, hogy mint kommu-

² Kurschus 2015: 53–70.

³ Lásd Gombrich 2003.

nikációs eszköz alkalmatlan volt arra, hogy egy hallótávolságon kívül lévő személlyel ténylegesen kommunikáljon a segítségével, arra azonban alkalmas volt, hogy a hallótávolságon belül lévők felé történő kommunikáció eszköze legyen. A sem jelen, sem távolban nem lévő adóssal folytatott beszélgetésen keresztül a sem a jelenben, sem távolban el nem érhető társadalmi státusz reprezentációjaként funkcionált.

Témánk szempontjából mindez egyáltalán nem elhanyagolható probléma. Amikor ugyanis feltesszük azt a kérdést, hogy a néprajztudomány területén „miképp járul hozzá a számítógépes kutatás megalapozásához a hagyományos kutatástervezés és -módszertan, illetve fordítva: hogyan alakítja át a gépi kutatás a humán feldolgozást; alakítja-e a tudományelméletet, és ha igen, hogyan”, akkor a legrövidebb és legtömörebb válasz gyanánt azt mondhatjuk, hogy nagyjából úgy, ahogyan a zsebszámológép hozzájárult a fenti történetben említett egyszeri férfi státuszreprezentációjához. A helyzet természetesen ennél bonyolultabb, megértéséhez azonban először magáról a néprajztudományról kell röviden szólnunk.

2.

Azt a kérdést, hogy mit is értünk néprajzon, azért érdemes felvetni, mert az elmúlt egy-két évtizedben a néprajztudomány Magyarországon is jelentős mértékben átalakult. Ennek az átalakulásnak egyértelmű jeleit fedezhetjük fel abban, hogy a hazai felsőoktatásban részt vevő néprajzi tanszékek közül Szeged, Pécs és Kolozsvár nevébe is felvette a néprajz mellé a kulturális antropológiát, s a többi intézmény képzésében is teret nyert az utóbbi diszciplína. A helyzet persze fordítva is megfogalmazható: mint néhány évvel ezelőtt Hofer Tamás egy konferencián⁴ rámutatott,

⁴ Hofer Tamás hozzászólása A komplex kultúrakutatás dilemmái II. konferencia egyik vitájában. A konferencia időpontja: 2003. október 16–17. Helyszíne: Miskolci Egyetem.

Magyarországon a kulturális antropológia a néprajz intézményi és akadémiai bázisán bontakozott ki. Ez a kapcsolat nyilvánvalóan nem volt konfliktusmentes, ám számunkra most a lényeg az, hogy a jelenlegi állapotra Hofer szerint az jellemző, hogy a néprajz bizonyos területeken – így a jelenkutatásban – antropológizálódott, az antropológia pedig érzékenyvé vált a néprajz problémafelvetéseire és történeti dimenziójára. Így ma már a néprajzos önmeghatározásként is gyakori és annak szinte szinonimájaként használt „etnográfus” szó jelentése egyre közelebb áll ahhoz, amit ezen az amerikai vagy brit gyakorlatban értenek. Az etnográfia így azt, az eltávolodás gesztusát magában foglaló, és döntően a terepmunkára – azon belül is a kutató résztvevő és megfigyelő perspektívát alkalmazó reflexiójára – épülő gyakorlatot jelenti,⁵ melyet az empirikus társadalomtudományok számos ága alkalmaz.

A néprajz alapvetően kvalitatív tudomány, többségében olyan, a kutató által a vizsgált társadalmi csoportban való jelenlétből, illetve a jelenségek megfigyeléséből származó, ún. „puha” adatokkal dolgozik, melyek konstrukciója jelentős mértékben függ a kutatói jelenlét minőségétől, vagyis a kutató reflexivitásától, megfigyelő és elemző kompetenciáitól. Nem kizárt, de nem is különösebben gyakori, hogy a „puha” adatok mellett „kemény” adatokkal, vagyis számszerűsített információkkal és kvantitatív módszerekkel dolgozik, s ha így is tesz, a feldolgozásra kerülő adatállományok ritkán érik el azt a méretet, mely egy humán ágens számára már átláthatatlan. Így azután érthető egyrészt az, hogy a néprajztudomány ritkán, leginkább múzeumi kontextusban épít adatbázisokat, másrészt pedig az, hogy a néprajzos egyetemi képzésekben elvétve találkozunk olyan kurzusokkal, melyek a tudományművelés informatikai hátterével és lehetőségeivel – például a kódolósos-visszanyeréses módszert alkalmazó

⁵ Így szerepel például a Royal Anthropological Institute „Education Outreach Programme” internetes honlapján is. Lásd: <https://www.discoveranthropology.org.uk/about-anthropology/fieldwork/ethnography.html>
Utolsó letöltés: 2018. március 25.

kvalitatív adatelemző szoftverek használatával – foglalkoznak, s ha vannak is ilyenek, azok leginkább speciális kollégiumok.⁶

A gyakorlatban a számítógép – és tágabban: a digitális eszközök – alkalmazása az elmúlt két évtizedben lényegében naiv, reflektálatlan módon váltotta fel az azonos vagy hasonló funkciójú analóg eszközöket. Így, elsősorban az intézményi modernizációs programok keretében az írógép helyére került az asztali számítógép, majd laptop. A filmes fényképezőgépeket és kamerákat lassan, döntően a költséghatékonyság és a jelentősen egyszerűbb képfeldolgozás okán felváltották a digitális eszközök. Az orsós, majd kazettás magnó helyett jött a minikazettás, majd digitális diktafon. A digitális eszközökre való átállás persze nem hagyta érintetlenül az etnográfia gyakorlatát: a terepmunkára menő néprajzkutató egyre inkább valamilyen elektronikus eszköz segítségével rögzíti jegyzeteit, és nem füzetekbe ír, interjúit valamilyen – többnyire a digitális diktafonhoz ingyen kapott – szoftver segítségével írja át ugyancsak valamilyen számítógépes szövegszerkesztő alkalmazás segítségével létrehozott dokumentumba. Eredményeit pedig egyre többször publikálja elektronikus formában az interneten – leginkább azért, mert forráshiány miatt a korábban papíron megjelenő társadalomtudományi folyóiratok lassan átálltak az internetes megjelenésre. Ezt a változást azonban egészen a közelmúltig⁷ nem kísérte a digitális eszközök haszná-

⁶ Jellemzően az elmúlt tíz évben átlagban két vagy három ilyen témájú szakdolgozat készült ezekben az intézményekben.

⁷ Két úttörő jelentőségű konferencia is zajlott ebben a témában a közelmúltban: egyrészt „Az etnográfia esélyei és lehetőségei a digitális térben” (Szeged, 2017. április 7–8., szervező: SZTE Kommunikáció- és Médiatudományi Tanszék és PTE Kommunikáció- és Médiatudományi Tanszék), másrészt a „Digitális társadalom- és médiatudomány: koncepciók, módszerek és gyakorlatok” című műhelytalálkozó (2017. szeptember 15–16-án, Pécsen, szervezők: PTE Kommunikáció- és Médiatudományi Tanszék, a Nyelv és kommunikáció doktori program és Az internetes forgalomirányító szolgáltatások szabályozási kérdései című OTKA-kutatás). A téma iránti fokozott érdeklődést jelzi, hogy a *Replika* folyóirat két száma is foglalkozott az elmúlt években a téma különböző aspektusaival (először

latával kapcsolatos elméleti és módszertani reflexió, vagyis nem vetődtek fel olyan alapvető kérdések, mint például az, hogy ha már egyszer hipertextként publikáljuk a szövegeinket, akkor mit tehet hozzá ez a sajátos, nem lineárisan építkező és különböző mediális természetű információhordozók integrációjára képes szövegforma az etnográfiai tudás termeléséhez és továbbadásához?

Ez a sajátos naivitás azonban nem csak a digitális eszközökkel kapcsolatos gyakorlatokra és nem is csak a magyar néprajztudományra jellemző. 1973-ban, amikor a fényképezőgép már meglehetősen hosszú ideje az etnográfusok, illetve kulturális antropológusok felszerelésének állandó része volt, és a fotográfiai készítése is szinte magától értetődően tartozott hozzá az etnográfiai kutatáshoz, egy amerikai kutató, Jay Ruby egy, az *Ethnographic Film Newsletter* című szaklapban megjelent cikkében ezzel a gyakorlattal kapcsolatban feltett két egyszerű kérdést: miért készítene az etnográfusok fényképeket, és miben különböznek ezek a nem etnográfusok által készített képektől? Az első kérdéssel kapcsolatban arra a következtetésre jutott, hogy maguk az etnográfusok sem tudnak egyértelmű választ adni rá. Mint írja: „amennyiben a képkészítés antropológiai tevékenység, elég észszerűnek látszik azt feltételeznünk, hogy fellelhető az antropológiai képkészítés tudományosan igazolható voltát bizonyító szakirodalom. Más szóval, ha az antropológusok időt, pénzt és

a 90–91. [2015/1–2.] szám „Belépés jelszóval! Online világok és kutatási módszerek” címmel, Lajos Veronika és Farkas Judit szerkesztésében; majd a 100. [2016/5.] szám „Tudományos megismerés és értelmezői gyakorlatok” tematikus blokkja Lajos Veronika szerkesztésében). Témánk szempontjából végül meg kell említenünk az MTA BTK Néprajztudományi Intézet két konferenciáját, melyek a terepmunka aktuális kérdéseit tematizálták. Az első, 2016 őszén megtartott belső műhelykonferencia az intézetben zajló néprajzi terepkutatásokkal foglalkozott. Ennek anyaga az intézet évkönyvének 2016. évi számában jelent meg. A második konferencia 2017 februárjában a magyar néprajztudomány más meghatározó intézményeinek előadóit is felvonultatta. Ennek előadásai az intézeti évkönyv 2017-es számában jelentek meg.

energiát fektetnek képek készítésébe, ezt azért kell tenniük, mert e médium egyedülálló minősége megteremti számukra annak lehetőségét, hogy rögzítsék, elemezzék és bemutassák a kultúra olyan vizuális megjelenési formáit, melyek másképp megfoghatatlanok, illetve ezen a módon ragadhatók meg legjobban. Azonban a dolog nyilvánvalóan nem így áll...”, hiszen efféle irodalmakra akkor Ruby nem nagyon bukkant. A második kérdésre az etnográfusok/kulturális antropológusok által készített képek elemzésével próbált meg választ találni, és arra jutott, hogy képek nem vagy alig különböznek azoktól, melyeket a hasonló helyeken és helyzetekben az egyszerű turisták készítenek.⁸

A néprajzi vagy etnográfiai, a szociográfiai és a fotóművészeti fényképkészítési és képhasználati stratégiák összehasonlítása alapján hasonló problémákat vázol fel egy 2004-es tanulmányában Bán András is, és a két világháború közötti példák elemzése kapcsán megjegyzi, hogy „a digitális technika új, de a bizonytalanságok régiek: a képesség kapcsán mostanában felbuzduló érdemi vizsgálatok körébe újra be kellene emelni a kép, a bizonyosság, a tudományosság és a társadalomvizsgálat egymásba hurkolódó kérdéseit is. Alap kutatásokra van szükség.”⁹

3.

Attól tehát, hogy a számítógép műveleti alkalmazása a néprajz tudományelméletének alakító tényezője legyen, még messze vagyunk. Vannak ugyanakkor olyan, elsősorban a kulturális antropológia és a szociológia területéről érkező impulzusok hatására kibontakozó, módszertanilag is innovatív egyéni kutatási kezdeményezések, melyek egyre erősebben hatnak vissza a néprajz elméleti diskurzusaira is. Ahogy nemzetközi, úgy hazai viszonylatban is ezen kutatások jelentős része az interneten vagy ott is szerveződő közösségek etnográfiai vizsgálatával foglalkozik.

⁸ Ruby 1973: 12.

⁹ Bán 2004: 17.

A cyberethnography vagy -anthropology¹⁰ területén az első hazai publikációt 2001-ben Gelléri Gábor jegyezte, aki saját internetes terepmunkájának példáján ismertette a cyberanthropology alapvető elméleti és módszertani megfontolásait,¹¹ azonban több mint tíz évnek kellett eltelnie ahhoz, hogy ez a kutatási irány – más, az etnográfiával tárgyi átfedésben is lévő – társadalom- és bölcsészettudományok után a néprajz olyan klasszikusnak tekinthető területein is megjelenjen és legitimé váljon, mint a folklorisztika.¹²

Ennek a fejleménynek a hátterében legalább két fontos tényezőt azonosíthatunk. Az egyik a különböző diszciplínák és cyber ágaik közti kapcsolat újragondolása volt. A cyber-társadalomtudományok kezdeti időszakában ugyanis a cyberközösségeket általában zárt entitásokként fogták fel, és így is kutatták. Ebben az esetben az alapvető kérdés az volt, hogy az adott diszciplína jellemző elméleti és módszertanai – például az etnográfia esetében a terep koncepciója és a résztvevő megfigyelés módszere – kiterjeszthetőek-e az új közegre, s ha igen, miképp. Később azonban, ahogy az internet a mindennapi élet egyre fontosabb részévé, pontosabban kiterjesztésévé vált, ez a szemlélet megváltozott. A cyber-közeg különössége és sajátossága sokkal kisebb súllyal van jelen a kutatásokban. A kutatók számára viszonylag gyorsan kiderült ugyanis, hogy a reális és az imaginatív vagy virtuális világok közötti erős különbségtétel nem igazán gyümölcsöző, s helyett a közösségek, identitások és kapcsolatok különböző

¹⁰ E tudományterület elnevezése körül meglehetősen nagy zűrzavar uralkodik, a cyber-, computer-, net-, online-, virtual- előtagok és az -ethnography, -anthropology utótagok különböző variációival találkozhatunk aszerint, hogy melyik kutató milyen irányból közelít a területhez és mit tart benne elsődlegesnek. A probléma részletesebb összefoglalását lásd Nagy 2015.

¹¹ Gelléri 2001.

¹² E területen lásd Domokos Mariann és Vargha Katalin magyar folklorisztikában úttörő jelentőségű kutatásait például Vargha 2005, Domokos 2005, Domokos 2013, Domokos–Vargha 2015, Vargha 2016, valamint Povedák 2007.

természetű kommunikatív terek közötti folyamatosságára kell koncentrálniuk.¹³ Ebből a szemléletváltásból pedig egyenesen következett a másik fontos tényező, vagyis az, hogy a kutatók az eredetileg nem cyberközösségekben vizsgált problémák kutatását egyre többször terjesztették ki a cyberközösségekre is.¹⁴

Az etnográfia digitális korszakba történő naiv átcsúszásának összefüggésében ezek a kutatások határozott előrelépést jelentenek, ugyanakkor azt is meg kell jegyeznünk, hogy elméleti és módszertani hozadékuk alapvetően a már meglévő koncepciók és gyakorlatok kiterjesztésére korlátozódik, vagyis elsősorban arra világítanak rá, hogy milyen módokon járul hozzá a számítógépes kutatás megalapozásához a hagyományos kutatástervezés és -módszertan. A következőkben három olyan kutatást mutatok be röviden, melyek ezen a szinten túllépve nemcsak azt a kérdést teszik fel, hogy miként boldogulhat az etnográfia a virtuális térreumok felé kitágított új világban, hanem azt is, hogy a digitális váltás milyen kihívásokat és lehetőségeket rejt az etnográfia számára, ezeknek miként tud eleget tenni, és hogy tudja ezeket kihasználni.

Borsos Balázs *A magyar népi kultúra regionális struktúrája a Magyar Néprajzi Atlasz számítógépes feldolgozása fényében* címmel 2011-ben megjelent munkájában a magyar népi kultúra térbeli tagoltságát bemutató, kilenckötetes Magyar Néprajzi Atlasz adatait dolgozta fel és elemezte számítógépes klaszteranalízis segítségével.¹⁵ A feldolgozott atlasz a teljes magyar nyelvterületet átfogó 417 kutatóponton ábrázol 627 jelenséget, a 19–20. század fordulójának állapotára visszatekintve. Az 1958-tól a 70-es évek elejéig tartó gyűjtést és a 159 gyűjtőt egy 12–15 fős törzsgárda

¹³ Wilson–Peterson 2002: 456–457.

¹⁴ Így például az irodalom-, kommunikáció- és vallástudomány határain mozgó Hubbes László Attila az apokaliptika tárgyában végzett történeti és recens kutatásait terjesztette ki az online közösségekre. Lásd Hubbes 2010; Hubbes 2011; Hubbes 2012.

¹⁵ Borsos Balázs DSc geológus, geomorfológus és etnográfus, az MTA BTK Néprajztudományi Intézet tudományos tanácsadója. Az itt bemutatott kötete Borsos 2011.

irányította, s az atlaszt 9 kötetben 1987–1992 között adták ki, magyarázó kötet nélkül. Borsos Balázs az elemzést az e célra az ő szakmai útmutatásai alapján készült EthnoMap Tools nevű szoftver segítségével egymaga végezte. Az atlasz összes adatának, mintegy 400 000 rekordnak a feltöltése és ellenőrzése három évig tartott. „A Néprajzi Atlasz térképlapjai önmagukban álló egységek: az egyes kultúrelemek térbeli elterjedését, fajtáját, létezését vagy éppen hiányát ábrázolják. Borsos Balázs viszont a számítógépes technika segítségével az egyes térképlapokat egymásra vetíti, és a több, mint 600 néprajzi jelenséget, »együttlátatja«. Mindezt kiegészíti az atlasz gyűjtőpontjainak összehasonlító statisztikai vizsgálatával, és a kialakult képet összeveti a közigazgatási, nyelvészeti, földrajzi és számos kulturális aspektus szerinti tagolódással. Ezzel a kultúra térbeli szerveződésének egzakt képét tudja felvázolni, képes a magyar nyelvterület nagy, közepes és kis kulturális régióinak körülhatárolására, lokalizálására, sőt, sajátos kultúrával rendelkező kis táji csoportok, mikrorégiók meghatározására is.”¹⁶ Borsos Balázs kutatása nagyon fontos eredményeket hozott, jelentős mértékben hozzájárult ahhoz, hogy a tér és kultúra kapcsolatát az eddigieknél¹⁷ részletesebben és pontosabban tudják körvonalazni és értelmezni a kutatók. A kutatási folyamat heroikus volta és a végeredmény médiuma azonban egyaránt jelzi a tudományterület alapvető problémáit: a kutatás eredményei mindeddig kizárólag könyv formában jelentek meg, jóllehet a narratív-elemző szövegeket, az adatokat feldolgozó és prezentáló adatbázis típusú közlésformát és a vizuális elemzés-áttekintést is lehetővé tevő adatvizualizációs megoldásokat azonos keretrendszerben, és egymásra vonatkoztatva elérhetővé tevő hipertext alapú kiadás az eredmények hasznosulását, a tudományos és köztudatba való beépülését jelentősen könnyítené.

Magyar és komplexitását, módszertani átgondoltságát tekintve nemzetközi viszonylatban is hasonlóan úttörő Vargyas Gábor *Brú fonó-foto-video-téka* címet viselő, 2008-ban megjelent hiper-

¹⁶ Balogh 2011.

¹⁷ Kósa–Filep 1978; Kósa 1990.

médiás projektje.¹⁸ Az interneten megtalálható anyag egy közép-vietnami hegyi törzs, a brúk között végzett két évtizedes terepmunka eredményeibe ad betekintést. A hipermediás keretrendszer lényegében a kutatás során készült álló- és mozgóképeket, rajzokat, hang- és filmfelvételeket közvetlenül a brú kultúra jelenségeit értelmező szövegekbe építi. Ennek a társadalomtudományi tudás létrehozása szempontjából két lényeges hozadéka van. Egyrészt a különböző mediális természetű információhordozókat egymással és az elemző szövegekkel is összekapcsolva sokkal komplexebbé teszi az interpretációt. Nemcsak azért, mert a hagyományos nyomtatott formánál lényegesen több vizuális információhordozót képes az értelmezés során mozgósítani, hanem azért is, mert az állóképek mellett mozgóképet és hangot is tartalmaz, melyek az értelmezés számára egészen más dimenziókat nyitnak meg, mint a szövegek. Másrészt ez a komplexitás nem csupán a szövegek felől értelmezhető. Az önmagukban egyedülálló, az értelmezésnek kiszolgáltatott terepmunkanyagokat a rendszer kontextualizálja. A képek például szekvenciákba rendeződnek, sorozatokat alkotnak, s így ott, ahol az egyedi kép értelmezési lehetőségei gyakorlatilag parttalanok, a szekvenciák által teremtett vizuális kontextus már lehatárolja, irányítja az interpretációt. Mindezt azonban a képszerűség sajátosságait – a lineáris, „történetmesélős” jelentésadás helyett az egyedi képet és a szekvenciát is összetett és nem hierarchizált utalásrendszerek totalitásaként – megtartva teszi lehetővé. Az egyes információhordozók ráadásul nem csak egy, hanem gyakorlatilag tetszőleges számú – és különböző mediális természetű más információhordozókat is tartalmazó – szekvenciának lehetnek részei, így a részben redundáns részben azon-

¹⁸ Vargyas Gábor DSc etnográfus, az MTA BTK Néprajztudományi Intézet tudományos tanácsadója, a PTE BTK Néprajz – Kulturális Antropológia Tanszék egyetemi tanára és doktori iskolájának vezetője. Az elemzett projekt – melynek létrehozásában a hipermediális keretrendszer tervezésével és megvalósításával magam is részt vehettem – elérhető az alábbi internet címen: <http://www.nti.btk.mta.hu/bffvt/index2.html> Utolsó leltetés: 2018. március 25.

ban mindig új elemeket is tartalmazó információknak köszönhetően egy kép vagy szekvencia értelmezése a rendszer bejárása során egyre komplexebbé is válik.¹⁹ A rendszernek ugyanakkor megvannak a maga korlátai. Egyrészt a szerző – bevallottan – lényegében csupán a publikáció új módjaként tekint a hipermédiára, s ennek az új módnak a lehetőségeit nem forgatja vissza az elemzés folyamatába. A szövegek elsősorban illusztrációként hivatkoznak a különböző audiovizuális anyagokra, ezek összetett hálózatát, egymásra utalásuk rendszerét tulajdonképpen maga a hipermédiális környezet struktúrája szinte automatikusan hozza létre. Másrészt az anyag publikálása óta eltelt időben a szerveroldali, illetve az internetböngészők által támogatott technológiák és szabványok változása miatt a statikus html²⁰ oldalakba épített

¹⁹ Ehhez lásd bővebben Nagy 2006.

²⁰ A HTML angol eredetű rövidítés (HyperText Markup Language, vagyis hipertextuális jelölőnyelv) egy leíró nyelv, melyet weboldalak készítéséhez fejlesztettek ki, és internetes szabvánnyá vált. Zömmel szöveges állományokként épül fel, egy sajátos szintaxis alapján, és tartalmazza azokat a kódokat, amelyek egy megjelenítő program számára leírják, hogy miként kell megjeleníteni, illetve feldolgozni az adott állomány tartalmát. Ez a kódrendszer különböző jellegű csatolók – ún. linkek – segítségével az adott dokumentum más dokumentumokkal való kapcsolatát is leírja. Amikor a hipertext nemcsak szöveges, hanem audiovizuális anyagokat is tartalmaz, vagy ilyenekkel kapcsolódik, hipermédiáról beszélünk. A hipertext/hipermédia legfontosabb sajátossága az, hogy a linkek és a különböző médiaanyagok összekapcsolásával a szövegek zárt rendje, s így a szerző által előre definiált olvasási-értelmezési szekvenciái felbomlanak vagy felbonthatóak, és így egy komplexebb olvasási stratégiát igényelnek, illetve hoznak létre. A hipertextet megalkotója, Ted Nelson olyan „nonszekvenciális írásnak” nevezi, mely „elágazik és a választás lehetőségét kínálja fel az olvasónak, s melyet a legjobban egy interaktív képernyő előtt lehet olvasni. Általában úgy szokták elgondolni, hogy szövegdarabok sorozatáról van szó, melyeket linkek kötnek össze, s így az olvasó különféle bejárési útvonalak között választhat” (idézi Landow 1996). Nelson koncepciójában a hipertext az ember asszociációkra épülő gondolkodását a lineárisan szervezett szövegeknél jobban tükrözi. A linkeket mint a szöveg mögötti más szövegekre utaló direkt utalásokat

különböző scriptek²¹ egy része nem vagy csak korlátozottan működik, így az oldal folyamatos frissítésre szorul. Végül szintén az alkalmazott technológia korlátaiból, az egyenként felépített statikus html oldalak alkalmazásából eredően, a rendszer bővítése akár új anyagokkal, akár új funkciókkal meglehetősen bonyolult és időigényes folyamat.

Végül a kutatás és prezentáció összekapcsolásának nagyon magas szintű megvalósításaként kell említést tennünk a Néprajzi Múzeum irányításával, „Kortárs jelenségek, mai tárgyak és mai emberek megismerésére és megismertetésére létrehozott múzeumi hálózat, a MaDok-program intézményeinek közös, lakóközsében utazó kiállításáról” a 2009-ben indult, majd 2011-ben megismételt *EtnoMobil*ról.²² A múzeum Frazon Zsófia²³ és Joó Emese²⁴ által koordinált projektje számunkra itt azért izgalmas, mert innovatív módon, a kiállított tárgyakon keresztül kapcsolja össze az offline/onsite tereket és az abban mozgó kiállítókocsit, az online kiállítást, illetve a blogot és a nyomtatott könyvet mint kommunikációs médiumokat. Az *EtnoMobil* kiállítás személyes hozzájárulásokra épült, a „nézők” maguk is alkotókká váltak, amikor saját személyes tárgyaikat vitték el a Mobilhoz, s az azokról készült digitális felvételekhez csatolták személyes történeteiket is, melyek azután így, együtt jelentek meg a virtuális térben, ahol a történetek továbbírhatókká váltak, hiszen a blog lehetővé tette, hogy az olvasók hozzászóljanak az egyes kép-szöveg egységekhez. A projekt így a kortárs tárgyhasználatok, s ezen keresztül a kultúra jelrendszerének használatába nyújtott betekintést.²⁵

tartalmazó hipertextet egyszerűen az intertextualitás explicit megvalósításának is szokás nevezni (Józsa é. n.). A kérdéshez lásd még Szűts 2001.

²¹ Az oldal működését szabályozó kisebb programok.

²² <http://www.neprajz.hu/madok/madok-projektek/etnomobil.html>
Utolsó letöltés: 2017. november 23.

²³ Frazon Zsófia PhD, etnográfus muzeológus, kurátor, a Néprajzi Múzeum tudományos titkára.

²⁴ Joó Emese etnográfus, múzeumpedagógus, közművelődési és kommunikációs szakember.

²⁵ A projekt leírását, elméleti háttérét lásd Frazon 2010.

4.

A három projekt rövid bemutatását azért tartottam fontosnak, mert jól jellemzik azt a problémahalmazt, mely körül a bevezetőben megfogalmazott kérdés második fele kikristályosodik, vagyis azt, hogy milyen hatással van a digitalizáció és a kutatás gépi eszköztára, illetve logikája a humán feldolgozásra és az etnográfia elméleteire. A következőkben ebből a körből három kérdést szeretnék röviden bemutatni, felvázolva tudománytörténeti hátterüket is:

1. az adatrendezés, strukturálás és integráció (valamint a multimedialitás) problémája;
2. a szerzőség, a szerző, az „adatközlő” vagy „bennszülött” és olvasó viszony átértékelésének, illetve a visszacsatolásnak a problémája;
3. a néprajz egyik központi kérdését jelentő terep problémája, és a tereptapasztalat komplexitásának értelmezése.

4.1. Margaret Mead, Gregory Bateson és a multimédia 1942-ben.

Az adatgyűjtés, -elemzés és a prezentáció területén alkalmazható technikai eszközök kérdése nagyon régen jelen van az etnográfia világában. Az etnográfusok szinte a kezdetektől használtak fényképezőgépet, de a társadalmi jelenségek rögzítésére, komplex elemzésére és bemutatására valószínűleg Franz Boas használta először az álló- és mozgóképek, hangfelvételek, valamint írott szövegek olyan integrált rendszerét, melyet ma multimédiának neveznénk.²⁶ Az ő hatása is érezhető tanítványa, Margaret Mead és Gregory Bateson 1936 és 1939 között végzett Bali szigeti kutatásán. Kettejük könyve, a *Balinese character* a vizuális antropológia meghatározó műve lett.²⁷ A fényképes és szöveges oldalak

²⁶ Ruby 1980; Nagy 2005.

²⁷ Bateson–Mead 1942.

párosaiból felépülő monográfiában – melynek lineáris szerkezetét a kereszthivatkozások bőséges rendszere lazítja fel – a képek nem pusztán illusztratív szerepet töltenek be, hanem az érvelés részét képezik,²⁸ a szerzők azonban ennél jóval többet akartak. Mint Bateson a könyv előszavában kifejti, egy olyan reprezentációs módszert szerettek volna létrehozni, melyben az állóképek és szövegek mellett a kutatás során különösen fontos szerepet játszó mozgóképfelvételek is megjelenhetnek: „arra voltunk kényszerítve, hogy nagyon takarékosan bánjunk a filmes nyersanyaggal, s így a bemutatás jövőbeli nehézségeit figyelmen kívül hagyva, azt terveztük, hogy az álló- és a mozgóképek együtt alkotják majd a viselkedéssel kapcsolatos felvételeink anyagát.”²⁹ Jóllehet a kor technikai adottságai között a bemutatás nehézségei megoldhatatlannak bizonyultak, és az alkalmazott módszer is számos problémát vetett fel,³⁰ a fényképek és filmek – illetve ezek szöveggel együtt történő – használatának, vagyis a multimedialitás és adatintegráció elméleti jelentőségére ők mutattak rá.

Mindketten úgy gondolták ugyanis, hogy a fényképek használatának egyik indoka az lehet, hogy azokat nem terhelik olyan kulturálisan rögzült asszociációk, melyekkel a szavak esetében meg kell küzdeni,³¹ vagyis a fényképek értelmezése független az értelmező kultúrájától. Ezt az elképzelést mára többszörösen

²⁸ Erről bővebben lásd Hagaman 1994.

²⁹ Bateson–Mead 1942: 50.

³⁰ Bateson az igazán fontos történéseknek akarta fenntartani a korlátozott mennyiségű filmes nyersanyagot, így a film és a fotó közti választás az adott pillanatban lényegesnek és lényegtelennek ítélt jelenségek közti kutatói megkülönböztetést implikálja. Bateson viszont éppen ezt akartak volna elkerülni azzal, hogy megfelelő technikai eszközökkel (a gyors fényképezést lehetővé tevő motoros filmtovábbítóval) és a fényképezés módszerének átalakításával (az esztétikai elvek mentén tervezett kompozíció helyett a gyors, az esemény lény lényegét középpontba állító és a lehető legtágabb kontextust megmutató képszerkezési mód alkalmazása) a lehető legtöbb fényképet készítsen egy megfigyelés során. A módszer következtében több mint 25 000 fotót készítettek, ami az átláthatatlanság határát súrolta.

³¹ Bateson–Mead 1942: xi–xii.

megcáfolták.³² Az a megfontolásuk azonban, hogy a megfelelően elkészített fényképek sorozatai segítségével a viselkedés illékony tintájával megjelenített kultúra precízebben írható le és tanulmányozható, mintha csak szöveggé konvertálnánk,³³ mára a vizuális antropológia egyik alapállítása lett. A leírás precizitása abból következik, hogy a fénykép viszonylag nagy hűséggel rögzíti a látvány egy-egy szeletét, s a rögzítés nézőpontjainak és a szeletek számának növelésével bizonyos mértékig ez a precizitás is növelhető. A fényképek azonban a leírás komplexitását is növelhetik. A szöveggel ellentétben ugyanis nem lineárisan, hanem szimultán módon, totalitásként reprezentálnak valamit, s ez a totalitás több egymást kiegészítő alternatív olvasatot tesz lehetővé. Egy kutatás összefüggésében az olvasatok közt ráadásul nemcsak a kutató, hanem a kutatók olvasatai is megjelenhetnek. Bateson éppen ezért a képeket és filmeket visszavitte a bennszülötteknek. A filmeket kézi vetítővel mutatta meg nekik, a falu lakói pedig kommentálták is a látottakat, például megbeszélték, hogy a lefilmezett szereplő tényleg transzban volt, vagy csak szimulált.³⁴ Bateson számára ezek a beszélgetések – amit egyébként interactive filmmaking eljárásnak nevezett – sokkal nagyobb jelentőséggel bírtak, mint maga a film, mert olyan dolgokat tudhatott meg általuk, melyeket kizárólag a filmek, képek vagy az interjúk útján nem tudhatott volna meg. Az etnográfiában ma széles körben elterjedt film vagy photo elicitation (vagyis film vagy fotó köré szervezett interjú) módszerét Batesonnak erre a gyakorlatára vezetik vissza.³⁵ Végül a szakirodalomban nem különösebben reflektált, ám témánk szempontjából különösen fontos az, hogy Bateson a könyv szerkezetével – az egy oldalra tett „kölcsonösen releváns” képekkel, a képek és szövegek közötti kereszthivatkozásokkal – alternatív olvasási utakat felkínálva, megpróbálja bevinni az olvasót is az etnográfiai értelmezés folyamatába. Ezzel

³² Például Worth–Adair 1977.

³³ Bateson–Mead 1942: XI–XII.

³⁴ Jackins 1988: 164–165.

³⁵ Lásd Krebs 1975.

egyrészt a Theodor Holme Nelson által majd negyven évvel később megfogalmazott hipertext gondolatát előlegezi meg, s nemcsak a prezentáció összefüggésében, hanem abban az értelemben is, hogy a hipertext „az adatintegráció és információközvetítés” egyik lehetséges módszere, ráadásul olyan módszer, mely „a tudást a hagyományosan lineáris szövegstruktúra helyett a kognitív emberi gondolkodást hívebben tükröző, nem lineáris, térbeli hálózatos rendszerben rendezi el”.³⁶

4.2. Collaborative ethnography és a web2

Bateson interactive filmmaking módszere előrevetíti a kutató – kutatott – olvasó szerep kiosztásának és felcserélésének azt a problémáját is, melyre az etnográfiaiban és általában a társadalomtudományi gyakorlatban az utóbbi időben egyre nagyobb szerepet játszó kollaboratív módszerek kínálnak megoldási lehetőségeket. Miután az etnográfia per definitionem kollaboratív tudomány, a szóhasználat némi magyarázatra szorul. A módszert kidolgozó Luke Eric Lassiter szerint³⁷ a kollaboratív etnográfia lényegében azt jelenti, hogy az együttműködés gyakorlatát a kutatási probléma kijelölésétől a terepmunkán át a tapasztalatainkat összegző írások elkészítéséig, az etnográfiai folyamat minden szintjére kiterjesztjük. Másképp szólva: az etnográfiai munka több stádiumára is jellemző együttműködéses módszereket vagy hajlandóságot a munkafolyamat háttéréből annak fókuszába állítjuk. Ez tehát nem egy új etnográfia jelent, hanem a használatos módszerek más megközelítését, hiszen maga a szó is szerepel a tudománytörténetben, csak míg korábban a – többnyire interdiszciplináris – kutatói együttműködésre vonatkozott, addig itt a lokális közösségekből származó munkatársakkal (local community consultants) történő valóságépítést jelöli.³⁸ A koncepciót

³⁶ Süethő 1999: 27.

³⁷ Lassiter 2005.

³⁸ Lassiter 2005: 17–18.

Lassiter a kognitív antropológiából vette, melyben arra a lokális szubjektumra,³⁹ akire az etnográfia korábban „adatközlőként” (informant, respondent) hivatkozott, egyre inkább „konzulens” vagy „szaktanácsadó” (consultant) megjelölést használták. A konzulens olyan „benszülött” gondolkodó társ, aki az etnográfussal együtt vállalkozik arra, hogy kidolgozza, illetve felfejtse a benszülött gondolkodás és kultúra kategóriáit, kapcsolatainak és jelentéseinek rendszerét.⁴⁰ Lassiter tehát úgy véli, az etnográfus nem a benszülött válla felett kukucskál,⁴¹ hanem mellette áll és vele egy irányba néz, s közben beszélget vele a látottakról. A collaborative ethnography módszerében nagyon fontos szerepe van az „egy irányba nézésnek”, vagyis annak, hogy az együttműködés már a terepmunka fókuszába kerülő társadalmi cselekvések kijelölésekor elkezdődik. Lassiter szerint nagyon fontos az, hogy a kutatók is érdekeltek legyenek a kutatás sikerében, ezt pedig úgy lehet elérni, ha az etnográfust foglalkoztató tudományos kérdéseket valós társadalmi problémák kezelésének összefüggésében kutatjuk. Az etnográfiai terepmunka célja az, hogy a kutató kommunikációs értelemben részesedjen⁴² azokban a felkészültségekben, melyek segítségével a lokális szubjektumok felismerik és kezelik a problémákat, s ezen keresztül újrateremtik, fenntartják a lokalitást. A részesedéshez azonban arra van szükség, hogy a kutató be is vonódjon a problémamegoldás diskurzusaiba, azok aktív résztvevője legyen, és hozzá tegye a saját felkészültségeiből a problémamegoldáshoz azt, amit tud – illetve amit a közösség adekvátan gondol. Az adekvát megoldás és hozzájárulás megtalálása a lokális tudásban való részesedés összefüggésében nehezen túlértékelhető kompetencia.

³⁹ A lokális szubjektum Arjun Appadurainál az az individuum, aki emlékezetében és testében a helyi, a lokalitás körülményeire alkalmazott és arra reflektáló tudás biztos hordozója. Lásd Appadurai 2001.

⁴⁰ Lassiter 2005: 25–48.

⁴¹ Geertz 2001.

⁴² A részesedés problémájához bővebben lásd Horányi 2001.

Azok a „beszélgetések”, melyek során az etnográfiai szövegek létrejönnek, legtöbb esetben a problémamegoldások színtereihez kapcsolódnak: értelmezik, dekonstruálják és újrateremtik azokat. A kutató és munkatársa együttműködésében létrehozott szövegeket minden esetben megvitatja – vagyis tovább írja – a közösség is.⁴³ A szövegek így ideális esetben a közösségi konszenzust tükrözik, aminek kétségtelen pozitívumai mellett azonban megvan az a veszélye is, hogy a kutatás a közösségi érdekérvényesítés eszközévé válhat, s így torzíthatja annak eredményét. Ez a helyzet azonban annak a lónak az „egyik” oldala, melynek másik oldalán azokat a stratégiákat találjuk, melyek segítségével a terepmunkát végző kutató hagyományosan megpróbálja uralni azokat a helyzeteket, melyekben a másokra vonatkozó tudását megszerzi. A kettő között nyeregben maradni nem egyszerű, a collaborative ethnography mégis erre tesz kísérletet azzal, hogy módszeresen rákérdezve a kutató státusára, az etnográfus dominanciájával jellemezhető, olykor konfliktusos szituációkkal szemben a kollaboráció, az együttműködés hangsúlyozásával a lokális tudás megértésének egy sokkal „dialektikusabb”⁴⁴ lehetőségét mutatja fel.

A szövegekkel kapcsolatos közösségi diskurzusok a szöveg és a kutatás befejezetlen karakterét jelzik, vagyis arra utalnak, hogy a szöveg maga is részévé válik a kutató és a kutatók közötti interaktív kapcsolatnak, beépül a diskurzusukba, és további reflexiókat tesz lehetővé. A kérdés csak az, hogy ezt a lehetőséget miként lehet kihasználni? Az etnográfiai kutatások az esetek többségében ugyanis lezárulnak az eredményeket – a kutatói értelmezéseket – bemutató szöveg(ek) könyvekben vagy folyóiratokban történő megjelenésével, s további diskurzusok legfeljebb azokban a szakmai körökben zajlanak, melyekhez ezek a médiumok eljutnak. Lassiter mindeddig sehol sem vetette fel a szövegek létrehozásával vagy közösségi értelmezésével kapcsolatban sem a hipermédia, sem a számítógépes mediáció, kivált az online

⁴³ Lasiter 2005: 140–144.

⁴⁴ Bornemann 2009: 238.

közösségi média használatának lehetőségét, ezek összekapcsolása azonban nagyon kézenfekvőnek tűnik. Egy interaktív hipermediális rendszer ugyanis – egyrészt a hipertext adatintegrációs lehetőségeinek köszönhetően, másrészt mint azt Batesonék esetében is láttuk, a képek segítségével létrehozható alternatív olvasatokon keresztül – a lokális szubjektumokkal való együttműködés „királyi útjaként” működhet. Ezen túl, a multimedialitás fentebb már elemzett lehetőségeit beépítve a rendszerbe, alkalmas arra is, hogy úgy legyen publikációs felület, hogy közben megőrzi a „végső szöveg” egységét. Egy közösségimédia-típusú környezettel egyesítve pedig a felhasználók által szabadon elhelyezhető hivatkozások és hozzáadott tartalmak (képek, szövegek stb.) segítségével arra is lehetőséget ad, hogy a szöveg „tovább éljen”. Végül fontos szempont lehet az is, hogy egy ilyen rendszer lehetőséget ad a közösségi konszenzust tükröző, reprezentatív-érdekérvényesítő szövegek mellett – épp ezek torzító hatását ellensúlyozandó, illetve ezek reprezentációs és érdekérvényesítési stratégiáit felfedő és értelmező – különvélemények, valamint az etnográfus értelmezésének együttes, egymással összefüggésbe állított megjelenítésére.

4.3. Multi sited ethnography és a hipertext

Az etnográfia most utolsóként kiemelt, állandó kérdése – melynek tekintetében az utóbbi időkben szintén jelentős változásokról beszélhetünk – a terep, vagyis az a kezdetben territoriális, ma már egyre inkább imaginárius értelemben vett „hely”, kommunikációs színtér, ahol az etnográfiai kutatás zajlik. E változások mögött az egyik fontos kérdés az, hogy miképpen tud lépést tartani az etnográfia a hagyományosan értelmezett illetékességi köréből, vagyis a „kis léptékű társadalmi keretektől” kilépő, majd a virtuális világokba belépő tárgyával, az emberrel, illetve hogyan tud módszertanilag alkalmazkodni azokhoz az egyre szélesedő – és egyre inkább elmosódó – globális horizontokhoz, melyek felé vizsgálatának tárgyai elindulnak. A problémára markáns és jelentős hatású választ adó George E. Marcus szerint a megoldás

az, ha az etnográfia/etnográfus követi a tárgyat, s a tárgy ki fogja jelölni, körül fogja írni a kutatás terepét.⁴⁵ E körülírás mibenlétét írásaiban személyek, tárgyak, metaforák, tervek, történetek, allegóriák, életutak, biográfiák és konfliktusok követésének összefüggésében tárgyalja. Marcus tudatában van annak, hogy a módszer az etnográfia határait feszegeti, tudatában van annak, hogy ezen az úton a terepmunka hatékonysága is csökkenhet, hiszen a klasszikus, egy-egy helyszínen hosszú időt eltöltő, állomásozó módszerrel szemben itt gyakran előfordulhat, hogy bár több színtéren, de az egyes színtereken rövidebb ideig van jelen a kutató egy-egy vizsgált közösségben, s így objektíve sincsen lehetősége a megismerés olyan fokát elérni, mint a klasszikus módszerekkel. Ugyanakkor – azon túl, hogy „klasszikus” problémáinak változásával is lépést tarthat – mégis fontosnak tartja a több színterű – multi sited – módszert, mert szerinte:

- a) A lehetősége éppen az, hogy a dolgoknak különböző lokális kontextusokban történő megjelenését követve nem csupán egyedi, hanem éppen az eltérő kulturális rendszerekben megváltozó jelentései tapinthatók ki. Ez pedig a kortárs etnográfia egyik fontos kutatási területe (gondoljunk például a népi kultúra elemeinek megjelenésére a posztmodern városi környezetekben).
- b) Másrészt így olyan témák, problémák kerülhetnek az etnográfia látóterébe, melyek addig illetékességi körén kívül álltak, ám amelyek tekintetében sajátos módszereinek – kivált a résztvevő megfigyelésnek vagy a vizuális etnográfának – mégis van relevanciája. E témák kutatása ugyanakkor azt is jelenti, hogy az etnográfia egy interdiszciplináris kutatási mező egyik résztvevőjévé válik.
- c) Végül Marcus szerint ez a lépés hozzájárul a tudományág elméleti horizontjának kitérítéséhez is. Utóbbi pedig kiemelkedően fontos számára, hiszen a kutatás tárgyának követése olykor olyan „helyszínekre” viheti az etnográfá-

⁴⁵ Marcus 1995: 108.

fust – például a média diskurzusai, vagy mint a módszert újabban tárgyaló szövegek megjegyzik, a cybertér –, melyekre elméletileg nincsen felkészülve.

A Marcus által felvetett gondolkodási irány így uralkodó tendenciaként jelenhet meg például a médiaantropológiában, kivált olyan területein, melyek az internettel foglalkoznak. Ezen a ponton válik igazán fontossá Marcus elmélete, mely a követés motívumának hangsúlyozásával éppen arra irányítja a figyelmet, hogy a terep nem olyan entitás, mely a kutatástól függetlenül is létezik. A terepre nem „kimegyünk”, hanem a terep a kutató és a kutatók kapcsolatának konstruktuma. Mint ilyen, egymástól nagyon különböző természetű kommunikatív színterekből jön létre – vagy másképp fogalmazva: számos ilyen színterre tagoldódik –, melyek közt az etnográfus problémafelismerése teremt kapcsolatot. A terepet tehát nem valamilyen territoriális elv tartja egyben, hanem a kutatás logikája.

5. Interaktív etnográfia

Összefoglalva az eddigieket: amennyiben tehát tudománytörténeti perspektívából közelítjük azt a bevezetőben feltett kérdést, hogy az etnográfia területén miképp járul hozzá a számítógépes kutatás megalapozásához a hagyományos kutatástervezés és -módszertan, illetve fordítva: hogyan alakítja át a gépi kutatás a humán feldolgozást és tudományelméletet, azt válaszolhatjuk, hogy a gépi intelligencia és az arra épülő eszközök használata szervesen illeszkedhet az etnográfia tradíciójába és problémátörténeteibe.

Az újmédia teoretikusaként ismert Lev Manovich 1997-ben a következőket írja: „ahogy a számítógép szerepe az eszköz felől egyre inkább eltolódik egy általános médiagép irányába, egyre többet »interfészselünk«, lépünk kapcsolatba alapvetően kulturális adatokkal: szövegekkel, fotókkal, filmekkel, zenékkal, virtuális közegekkel. Vagyis most már nem a számítógéppel lépünk kapcsolatba, hanem digitális formában kódolt kultúrával. Sze-

retném bevezetni a »kulturális interface« fogalmát, amely azokat a kialakulóban lévő interface-eket írná le, melyeket a webhelyek, CD-ROM-ok, DVD-ROM-ok, multimédia enciklopédiák, online múzeumok, számítógépes játékok és más kulturális termékek tervezői használnak.⁴⁶ A kulturális interfész tehát olyan technikák és eszközök korpuszát jelenti, melyek segítségével a számítógépen mint általános médiumon keresztül bármilyen kulturális adattal – illetve a digitalizált kultúra elemeivel – interakcióba léphetünk. Ez a meghatározás nem áll távol az etnográfától, mint olyan tevékenységtől, melynek értelme – Clifford Geertz szerint – nem az, hogy választ adjon az emberek kérdéseire, hanem az, hogy megmutassa, egy-egy kérdésre a különböző kultúrákban milyen válaszok születtek.⁴⁷ Az etnográfia tehát egyrészt tudományos gyakorlatként, másrészt a globalizáció nyomán, az – itt nem szitokszóként értelmezett – multikulturális környezetekben egyre népszerűbb megismerő tevékenységként sajátos, kulturális interfész szerepet tölthet be. Olyan interfész lehet, melynek kialakításában – úgy az interfészeknél elengedhetetlen érintkezési felületek megkonstruálásában, mint a kódok meghatározásában – közösen vesznek részt azok, akiket ezek összekötnek. Az etnográfia, mint kulturális interfész, kommunikációs értelemben így felkészültséggé válhat egyes társadalmi problémamegoldó folyamatokban, illetve maga is hasonló felkészültségek közös és kölcsönös elérhetővé tételének színtere lehet. A következőkben ennek egy lehetséges megoldását szeretném röviden bemutatni egy saját kutatási projekttem példáján.

Azt a kutatási módszert, melyet az eddig bemutatott problémák mentén az elmúlt években végzett kutatásaim során kidolgoztam, interaktív etnográfának nevezem. Ez egy olyan társadalomtudományi kutatási módszer, mely az interaktivitás fogalmát használja az etnográfiai megismerési folyamat leírására, illetve szervezésére, és ezzel szoros összefüggésben a kutatás tervezésétől a publikációig, illetve azon keresztül az olvasókkal történő kom-

⁴⁶ Manovich 2001: 35.

⁴⁷ Geertz 2001.

munikációig tudatosan, tudományos eszközként alkalmazza az interaktív hipermediális technikákat. Utóbbi révén az interaktív etnográfia mint produktum, egy etnográfiai hipermediális hálózati elemző-környezetként (EHHE) működik. Az EHHE az etnográfiában különböző elméleti irányokból megfogalmazódó, egyre intenzívebben használt – és az előzőekben már részletezett – többterepű etnográfia és a kollaboratív módszerek tanulságait ötvözi a közösségi tudástermelés hálózati szoftveres technológiáival egy olyan hipermediális keretrendszerben, mely

- egyrészt a kutatás teljes folyamatát összefogja a tervezéstől a terepmunkaanyagok rögzítésén és kódolásán keresztül a publikációig és a kutatott közösséggel, a kutatókkal, illetve az olvasókkal történő kapcsolattartásig;
- másrészt az etnográfiában többnyire uralkodó szövegcentrikus kutatási módszereket a vizuális antropológia módszereinek alkalmazásával tágítja, s így az interpretációt komplexebbé teszi;
- végül a kutatás során keletkezett, illetve az interpretációba bevont, különböző mediális természetű információhordozókat egy közös rendszerben integrálja, és azonos interfészen teszi elérhetővé.

Ahogy a kollaboratív jelző, úgy az interaktivitás esetében is indoklásra szorul a szóhasználat, hiszen az interaktivitás a szó széles körben elterjedt értelmezése alapján egyrészt feltételezi a kollaborativitást és viszont, másrészt az etnográfia magától értetődő módon interaktív. Az indoklásra keresztül ráadásul arra is rá tudunk világítani, hogy miként függ össze az EHHE-tanulmányom alapvető kérdésével, vagyis azzal, hogy milyen hatással lehet a digitalizáció és a számítástechnikai eszközök használata az etnográfia elméleteire és módszereire. Ezt a kérdést a következőkben egy, az etnográfia szempontjából különösen is aktuális és alapvető kérdés, a terep és a terepmunka fentebb Marcus kapcsán már bemutatott változásainak összefüggésébe állítom.

Ahogy korábban láttuk, az interaktivitás fogalma már Batesonnál is jelen volt, műveleti fogalommá azonban csak az utób-

bi húsz évben vált. Jóllehet használatával ritkán találkozunk, az általa hivatkozott tartalom mindig is fontos volt az etnográfia-ban, hiszen az etnográfus terepre vonatkozó tudása azokból az interakciókból táplálkozik, melyeket meg tud figyelni, illetve amelyek közegébe bele tud bonyolódni, s folyamatosan arra törekszik, hogy válaszra bírja ezt a közeget. A „közeg válaszol” kifejezés Szűcs Zoltántól származik, aki szerint az interaktív „kifejezés leginkább a multimédiára használható, ahol egy bizonyos képre vagy szövegre kattintva a közeg válaszol; egy klippel, képpel, vagy a hypertext esetében újabb szöveggel”.⁴⁸ Ez egy szűkebb, technikaközpontú megfogalmazás, mely ráadásul mindig magában hordozza a tervezettség koncepcióját is, hiszen az ilyen technikai közegek válasza általában előre meghatározott vagy legalábbis programozott. Tágabban, kommunikációelméleti összefüggésben⁴⁹ az interaktivitás a kommunikációs folyamatok értelmezésének egyik lehetséges perspektívája, mely arra koncentrál, hogy a kommunikatív aktusok hozadéka az egymást követő aktusokból álló folyamatban részt vevő ágensek számára a folyamat (például egy párbeszéd) megértése szempontjából felkészültségé,⁵⁰ vagyis a megértés szempontjából szükséges tudássá, illetve

⁴⁸ Szűcs 2001: 35.html.

⁴⁹ Például Rafaeli 1988.

⁵⁰ A felkészültség fogalmáról kommunikációelméleti összefüggésben többnyire tudásként, vagy a tudás analógiája szerint szoktunk beszélni. „A gyakorlati tudás kapcsán jártasságokat, míg az elméleti tudás kapcsán kompetenciát, vagyis egy felhalmozott ismerethalmazt érthetünk a felkészültségen” (Pete–Fazekas 2008). Antropológiai összefüggésben azonban hangsúlyozni kell a felkészültség fogalmának összetettségét és inerperszonális karakterét. Összetett, amennyiben „a felkészültségek részben tudásokból, részben hiedelmekből állnak, valamint különféle attitűdökből, például érzelmekből, szándékokból és így tovább” (Horányi 2015) és interperszonális, amennyiben egykommunikációs helyzetben bármi felkészültséggé válhat, azonban bármilyen felkészültségnek a felkészültség-értéke, vagyis az, hogy egy probléma felismerésének vagy megoldásának összefüggésében relevánsnak bizonyul-e, a kommunikációba bevonódott ágensek kapcsolatának függvénye.

attitűddé válik. Bizonyos esetekben, ahol a színtér részleges rögzítésére tesznek kísérletet, magába a színtérbe is beépül (például hosszú viták egy nyomtatott folyóiratban). Az utóbbi megfogalmazás az előbbinek is keretet ad, hiszen a hipertext vagy hipermedia nem más, mint egy ilyen, hosszabb-rövidebb ideig, többé-kevésbé rögzített kommunikációs színtér, mely sok esetben egy valós időben lezajlott kommunikációs folyamatot rögzít.

Az EHHE összefüggésében ilyen értelemben beszélhetünk az interaktivitásról, és ennek kapcsán lehet talán a legjobban kifejezni azt is, mit jelent az, hogy az interaktivitás az etnográfiai megismerési folyamat leírására, illetve szervezésére is érvényes. Korábban hivatkoztam George E. Marcusra, aki a multi-sited ethnography módszerét nagyon leegyszerűsítve úgy határozza meg, hogy az etnográfus a kutatás tárgya után megy, azt követi, s a tárgy – a követés folyamatán keresztül – körülírja a kutatás terepét.⁵¹ Marcus gondolatmenetében nem annyira a „követés” kifejezés esetleges metaforikus használata az érdekes – bár ez összecseng a hipertext alapvető metaforáival, s a koncepció is könnyen átfordítható egy hiperlink hálózatra –, hanem az, hogy maga a terep fogalma válik metaforává. Megközelítésében a terep nem egy topográfiai értelemben kijelölhető hely, hanem egy olyan koncepció, olyan „gondolati tér”, mely összekötni azokat a site-okat, vagyis – nagyon különböző – kommunikációs szintereket, melyeken a kutató a kutatás logikáját meghatározó problémát követi; illetve amely éppen ezekből a site-okból interaktív módon jön létre. A terep ilyen elgondolásában az egyes site-ok a kutató számára olyan felkészültségeket jelentenek, melyek mentén a terep, amit a kutató – a kutatás összefüggésében – értelmezett problémafelismerő és -megoldó színtere kifeszül. A kutató ugyanis ezeken a – „bennszülöttekkel”, lokális szubjektumokkal vagy helyi asszisztensekkel – kollaboratív módon létrehozott site-okon, mint kommunikációs szintereken, az ezeken folyó interakciók révén részesedik azokban a tudásokban, attitűdökben

⁵¹ Marcus 1995: 108.

stb., melyek egyrészt az adott színtérhez kapcsolódó probléma, másrészt épp ezeken keresztül a kutatás problémájának megoldásában számára felkészültségekké válhatnak. Kommunikációelméleti összefüggésben ugyanis a kommunikáció színtere a részt vevő ágensek adott problémafelismerése, illetve -megoldása során mozgósított és egymás számára kölcsönösen elérhetővé tett felkészültségei mentén jön létre. A színtérben a „tér” itt is metafora, s bár adott esetben utalhat egy konkrét térre, ám még akkor is a kommunikatív eseménynek azt a gondolati terét jelenti, „ahol” egy problémakezelésben részt vevő ágensek a probléma felismeréséhez és/vagy megoldásához általuk szükségesnek vélt felkészültségeiket egymás számára kölcsönösen elérhetővé teszik. A terep ebben a megközelítésben egy roppant összetett, rétegzett és sajátos történetiséggel rendelkező színtér, mely magában foglalja az egyes – önmagukban is problémakezelő színtérként értelmezhető – site-okat, s az ezekből például tapasztalatként reflektálva felépített konstrukciót csakúgy, mint azokat a színtereket, melyeken a kutató más problémakezelési folyamatok összefüggésében erre a konstrukcióra (is) hivatkozik. Az EHHE pedig lényegében nem más, mint ezeknek a színtereknek az újjáépítése, re-konstrukciója.

Az újjáépítés alapja természetesen a kutatói jelenlét a terepen, melynek során a részvétel és megfigyelés – a kollaboratív és vizuális etnográfia elvei segítségével kiterjesztett – hagyományos módszereivel rögzíti a tapasztalatait. A rögzítés elsősorban digitális eszközökkel történik, keretét pedig már az EHHE adja, s a kutató ezen a rendszeren keresztül vonja be a helyi munkatársait is az együttműködés folyamatába. A cél természetesen nem a kutatási site-ok virtuális tükrözése, egyfajta „hologram” létrehozása, hanem a site-hoz kapcsolódó kommunikációs színtér rekonstrukciója az adott színterre jellemző felkészültségek, illetve kapcsolataik explicitté tétele, és azok reflexív, értelmező (sűrű) leírása segítségével az EHHE-ben. Az explicitté tétel a felkészültségek különböző (szöveges, álló- és mozgóképes, illetve hangképes) mediális természetű információhordozók segítségével történő rögzítéséből, és a szó tág értelmében vett (vagyis a

szövegek mellett pl. képekre épülő) „adatolásából” áll, s ezekre épülhet elemző leírásuk.

Az EHHE hátterét egy módosított MediaWiki szoftver szolgáltatja, mely számos lehetőséget ad nem csupán az együttműködésbe bevont munkatársak részvételére, de a színterek rekonstrukciójának minél összetettebb megvalósítására is. A MediaWikiben a tartalmi elemek egy egységes címhez tartozó lapok komplexeként jelennek meg. A színteret alkotó egyes felkészültségelemek önálló lapokként rögzíthetjük. A lap főszövege alapesetben a kutatásban részt vevők egyéni meglátásait tükrözi, vagyis például azt, ahogyan a kutató rögzít egy eseményt. A főszöveghez kapcsolódó, ugyanazon cím alatt megjelenő vitalapon azonban bármelyik munkatárs kifejtheti véleményét a főszöveg narratívájáról, és konszenzus esetén kiegészítése be is kerülhet a főszövegbe. A szoftver azonban rögzíti ezeket a változásokat, és bármelyik korábbi állapotot megtekinthetővé vagy visszaállíthatóvá teszi.

Az egyes címek alá tartozó lapkomplexeket a rendszer többféle módon teszi összefűzhetővé. A lapkomplexek egyrészt kategóriába rendezhetőek. A kategóriákhoz a szoftver automatikusan létrehoz egy olyan lapot, melyen listázza az adott kategóriához tartozó lapkomplexeket (egy lapkomplex több kategóriához is tartozhat), és a kategórialap a továbbiakban maga is lapkomplekxként működik. Amennyiben az egyes lapkomplexek egy adott színtérhez tartozó felkészültséget reprezentálnak, a kategórialap a színtér reprezentációjának tekinthető. Ugyanolyan virtuális módon létezik, mint maga színtér, és ugyanúgy azok a felkészültségek konstituálják, mint a színteret. A lapkomplexek közti kapcsolat létrehozásának legegyszerűbb módját a linkek jelentik. A linkek egyedi lapkomplexeket kötnek össze, ezzel különböző logikai kapcsolatok kifejezését és az egyedi lapkomplexekből összeállított, különböző struktúrájú szekvenciák építését teszik lehetővé, így pedig az értelmezés eszközeként használhatóak.⁵²

⁵² A témához lásd Trigg 1983; Józsa [é. n.]: szakd6.htm; Bernstein 1998.

A linkek és linkhálózatok alapértelmezés szerint nem rendelkeznek olyan saját, leíró lappal, mint a kategóriák, de létre lehet hozni ezekhez kapcsolódva is hasonlókat. Az egyszerű linkeknél összetettebb, és szintén önálló leíró lappal rendelkező interpretációs eszköz az ún. WikiSzavak rendszere. A WikiSzó – a kifejezés írásképe is mutatja – két szó összetételéből áll, melyeket nagy kezdőbetűvel, szóköz nélkül egybeírunk. Az ilyen szavak a kategóriákhoz nagyon hasonlóan működnek, csak míg azokat külön hivatkozással kell a szöveg végére illeszteni, addig a WikiSzavak létrehozását gépeléskor automatikusan érzékeli a rendszer, és a szóhoz egy linkkel kapcsolja a szintén automatikusan létrehozott leíró lapot. A WikiSzavak kulcsszóként vagy kódként alkalmazhatóak, így a kódolósos-visszanyeréses tartalomelemző eljárás céljaira is használhatóak.⁵³ A kategóriák is és a WikiSzavak is – különböző módon, de egyaránt – egymásba ágyazhatóak, így az elemzés során mindkettőből struktúra építhető.

Az EHHE működése sok tekintetben tükrözi a megalapozott⁵⁴ elmélet (grounded theory) kutatási módszerét. Építését ugyanis bárhol el lehet kezdeni, nem igényel egy előzetesen létrehozott struktúrát és elméleti konstrukciót, hanem a struktúra egyetlen lapból – vagyis egyetlen tapasztalattal kapcsolatos reflexióból – is felépíthető, és a tapasztalatok függvényében változik, illetve változtatható. A struktúra felépülésében ezenfelül megnyilvánul valamennyi automatizmus is: a kódokból, kategóriákból, linkekből részben automatikusan létrejövő elemzési szint azonban a kutató számára rögtön lehetőséget ad az elméletalkotásra és az elmélet revíziójára.

⁵³ Jóllehet ezen a területen a MediaWiki alapértelmezett képességei korlátozottak.

⁵⁴ A megalapozott elmélet módszere az empiria és az elméletalkotás folyamatos, körkörös visszacsatolására építő rendszere. Bővebben lásd Glaser–Strauss 1967.

6.

Úgy vélem, hogy a leírtakból a kifejtés minden töredezettsége és korlátozott volta ellenére is egyértelműen látható, hogy az etnográfia elméletének és módszertanának több olyan eleme is van, melyekhez a számítástechnika, a digitalizáció érdemi módon kapcsolódhat. Az új technológiák használata, beépítése ezekbe az elméletekbe és módszerekbe fejlesztésüknek logikus lépése lehetne. Jelenleg azonban a tudományág képviselői és intézményei többségének viszonya az új technológiákhoz inkább reflektálatlansággal jellemezhető. Az infrastrukturális fejlesztési pályázatok állandó tétele a számítástechnikai eszközök beszerzése, az új, korszerű eszközök mellől azonban sokszor hiányzik az alkalmazásukra vonatkozó korszerű tudás, és a kihasználásukat lehetővé tevő innováció. Megmaradnak tehát a modern enteriőrök reprezentatív kellékeinek, méregdrága levélnehezékeknek, s funkciójuk nagyjából annyi, mint azé a bizonyos telefont helyettesítő zsebszámológépé.

Ahhoz, hogy a helyzet megváltozzon legalább két területen van szükség előrelépésre. Az egyik terület a diszciplína belső ügye: szükség lenne arra, hogy a problémakörhöz kapcsolódó elméleti, módszertani írások és a gyakorlatot bemutató esettanulmányok akár önálló felületen is megjelenhessenek a szakmai nyilvánosságban. Az is fontos lenne, hogy a néprajzi, kulturális antropológiai képzésekben megjelenjen vagy hangsúlyosabban legyen jelen ez a tematika, és ne csak úgy tekintsenek rá, mint publikáció- vagy prezentációtechnikai ismeretre, hanem elsősorban úgy, mint egy olyan komplex eszközkészletre, mely befolyásolja s olykor meg is határozza az etnográfiai gondolkodást. A néprajzi írás, a film vagy fotográfia esetében már végbement egy ilyen fordulat, ideje, hogy a digitalizáció területén is megtörténjen.

A másik terület azonban az etnográfia társadalmi beágyazottságával, „társadalmi hasznosságával”, az etnográfiai módszerekkel megszerezhető tudás alkalmazhatóságával kapcsolatos társadalmi sztereotípiák kérdése. Ahogy láttuk, a kollaboratív etnográfiai kutatások és a számítógépes mediációra építő, közösségi médiát használó, bevonódást feltételező módszerek többnyire arra építe-

nek, hogy a kutatásnak, melynek megvalósításában alkalmazásra kerülnek, „tétje van”. Sokszor olyan alkalmazott – például a public anthropology körébe tartozó – kutatásokról van szó, melyek kimondott célja a különböző szintű politikai vagy gazdasági döntések előkészítése vagy befolyásolása, a stratégiaalkotás vagy a közösségi érdekérvényesítés. Jelenleg azonban az etnográfiai kutatások segítségével megszerezhető tudás – akár a néprajz, akár a kulturális antropológia felől közelítjük – gyakorlati jelentősége Magyarországon meglehetősen csekély. Passzív tudás. Úgy az olvasóközönség, mint a különböző szervezetek, intézmények és döntéshozók számára, kevés kivétellel a civil, de kivált a kormányzati szférában nem bír különösebb értékkel. Ebből következően pedig meglehetősen nehéz például egy kollaboratív etnográfiai kutatásban érdekeltté tenni a helyi közösség tagjait, hiszen nem érzik annak súlyát, tét nélkülinek gondolják, mert hiába vesznek részt benne, a felhalmozott tudásuk, tapasztalatuk hasznosulására nem látnak esélyt, nem lévén előttük pozitív példa. A két terület szorosan összetartozik tehát: ahhoz, hogy a számítástechnika, a digitalizáció reprezentatív helyett eszköz szerepbe lépjen, eszközként értékkel kell felruházni. Ez a diszciplína képviselői előtt álló feladat.

Irodalom

- Appadurai, Arjun 2001. A lokalitás teremtése. *Regio* 12 (3): 3–31.
- Balogh Balázs 2011. Borsos Balázs: A magyar népi kultúra regionális struktúrája 1–2. In Borsos Balázs: *A magyar népi kultúra regionális struktúrája 1–2*. Budapest: MTA Néprajzi Kutatóintézete. Borító.
- Bán András 2004. Magyar képek. In Biczó Gábor (szerk.): *Vagabundus. Gulyás Gyula tiszteletére*. Miskolc: Kulturális és Vizuális Antropológia Tanszék. 17–22.
- Bateson, Gregory – Mead, Margaret 1942. *Balinese Character: A Photographic Analysis*. New York: New York Academy of Sciences.
- Bernstein, Mark 1998. Patterns Of Hypertext. In Shipman, Frank – Mylonas, Elli – Groenback, Kaj (eds.): *Proceedings of Hypertext '98*.

- New York: Association for Computing Machinery. URL: <http://www.eastgate.com/patterns> Utolsó letöltés: 2018. március 25.
- Bornemann, John 2009. Fieldwork Experience, Collaboration, and Interlocution. The „Metaphysics of Presence” in Encounters with Syrian Mukhabarat. In John Borneman – Abdellah Hammou-di (eds.): *Being There: The Fieldwork Encounter and the Making of Truth*. Berkeley: University of California Press. 237–258.
- Borsos Balázs 2011. *A magyar népi kultúra regionális struktúrája 1–2*. Budapest: MTA Néprajzi Kutatóintézete.
- Budka, Philipp – Kremser, Manfred 2004. CyberAnthropology – Anthropology of CyberCulture. In Khittel, Stefan – Plankensteiner, Barbara – Six-Hohenbalken, Maria (eds.): *Contemporary issues in socio-cultural anthropology: Perspectives and research activities from Austria*. Vienna: Loecker. 213–226.
- Domokos Mariann 2005. Népi írásbeliség a technika korában, avagy az sms mint folklórszöveg. In Gulyás Judit – Tóth Arnold (szerk.): *Mindenes Gyűjtemény II. Tanulmányok Küllös Imola 60. születésnapjára*. Budapest: ELTE BTK Néprajz Tanszék. 289–312.
- Domokos Mariann 2013. Az elektronikus folklór gyűjtéséről. In Berta Péter – Ispán Ágota – Magyar Zoltán – Szemerényi Ágnes (szerk.): *Ethno-Lore. A Magyar Tudományos Akadémia Néprajzi Kutatóintézetének évkönyve. XXX*. Budapest: Akadémiai Kiadó. 292–320.
- Domokos Mariann – Vargha Katalin 2015. Elektronikus választási folklór 2014. *Replika* 90–91: 141–169.
- Escobar, Arturo et al. (eds.) 1994. Welcome to Cyberia: Notes on the Anthropology of Cyberculture [and Comments and Reply]. *Current Anthropology* 35 (3): 211–231.
- Frazon Zsófia (szerk.) 2010. *EtnoMobil. Mozgásban a kortárs kultúra. /MaDok-füzetek 7/* Budapest: Néprajzi Múzeum.
- Geertz, Clifford 2001. Sűrű leírás. Út a kultúra értelmező elméletéhez. In *Az értelmezés hatalma. Antropológiai írások*. Budapest: Osiris Kiadó. 194–226.
- Gelléri Gábor 2001. Mit keres az antropológia a cyberben – és a cyber az antropológiában? *Tabula* 4 (2): 270–286.
- Glaser, Barney G. – Strauss, Anselm L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing.

- Gombrich, Ernst H. 2003. Elmélkedés egy vesszőparipáról, avagy a művészi forma gyökerei. In Horányi Özséb (szerk.): *A sokarcú kép. Válogatott tanulmányok a képek logikájáról*. Budapest: Typotex. 23–40.
- Hagaman, Dianne 1994. Connecting Cultures. Balinese Character and the Computer. *The Sociological Review* 42 (1): 85–102.
- Hine, Christine 2000. *Virtual ethnography*. London, Thousand Oaks, New Delhi: Sage.
- Horányi Özséb 2001. A kommunikációról. In Béres István – Horányi Özséb (szerk.): *Társadalmi kommunikáció*. Budapest: Osiris Kiadó. 22–34.
- Horányi Özséb 2015. *A kommunikáció mint állapot. Rövid áttekintés. Kézirat*. URL: http://www.ozseb.horanyi.hu/kozelet/tanulmanyok/a_kommunikacio_mint_allapot_150607.htm Utolsó letöltés 2018. március 25.
- Hubbes László 2010. Debating Contemporary Internet-Based Apocalyptic Discourse in Class. In: Bakó R. – Horváth G. (eds.): *Argumentation. Proceedings of the first international conference on teaching Argumentation and Rhetoric*. Cluj-Napoca: Transylvanian Museum Society. 46–59.
- Hubbes László 2011. Új magyar eszkatológia: apokaliptizáló retorika a világhálón – Néhány magyar nyelvű összeesküvés-szemlélet alapú honlap retorikai és szemiotikai tartalomelemzése. In Balogh F. András – Berszán István – Gábor Csilla (szerk.): *Újrateremtett világok. Írások Cs. Gyimesi Éva emlékére és életművének bibliográfiája*. Budapest – Kolozsvár: Argumentum Kiadó. 413–420.
- Hubbes László (szerk.) 2012. Etno-Pogányok: retorikai fogások a hálón innen és túl. Kolozsvár: Erdélyi Múzeum Egyesület. 119. URL: http://issuu.com/maybe/docs/etno_poganyok. Utolsó letöltés: 2018. március 25.
- Jacknis, Ira 1988. Margaret Mead and Gregory Bateson in Bali. Their Use of Photography and Film. *Cultural Anthropology* 3 (2): 160–177.
- Józsa Péter [é. n.]. *Irodalom a digitális közegben v1.0*. URL: <http://mek.niif.hu/02300/02313/html/index.htm> Utolsó letöltés: 2018. március 25.
- Kósa László – Filep Antal 1978. *A magyar nép táji-történeti tagolódása*. Budapest: Akadémiai Kiadó.

- Kósa László 1990. *A paraszti polgárosodás és a népi kultúra táji megoszlása Magyarországon (1880–1920)*. Debrecen: Kossuth Lajos Tudományegyetem.
- Krebs, Stephanie 1975. The Film Elicitation Technique. In Hockings, Paul (szerk.): *Principles of Visual Anthropology*. The Hague: Mouton. 283–301.
- Kurschus, Stephanie 2015. *European Book Cultures: Diversity as a Challenge*. Wiesbaden: Springer VS.
- Landow, George P. 1996. Hypertextuális Derrida, posztstrukturalista Nelson? In Sugár János (szerk.): *HYPertext + MULTIMEDIA*. Budapest: Artpool. URL: <http://www.artpool.hu/hypermedia/landow.html> Utolsó letöltés: 2018. március 25.
- Lassiter, Luke Eric 2005. *The Chicago Guide to Collaborative Ethnography*. Chicago: The University of Chicago Press.
- Manovich, Lev 2001. A film mint kulturális interface. *Metropolis* 7 (2): 24–43.
- Marcus, George E. 1995. Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography. *Annual Review of Anthropology* (24): 95–117.
- Nagy Károly Zsolt 2005. A fotóantropológia és az antropológiai fotó lehetőségei a „digitális képi forradalom” korában. *Tabula* 8 (1): 91–110.
- Nagy Károly Zsolt 2006. Interaktív etnográfia. *Tabula* 9 (2): 333–350.
- Nagy Károly Zsolt 2015. Ösvény a dzsungelben. *Replika* 90–91: 39–56.
- Pete Krisztián – Fazekas Piroska 2008. Felkészültség. In Horányi Özséb (főszerk.): *Kommunikációtudományi Nyitott Enciklopédia*. PTE Kommunikáció Doktori Program. Ad vocem.
- Povedák István 2007. A média szerepe napjaink populáris kultúrájában. In Ambrus Vilmos – Schwarz Gyöngyi (szerk.): *Változó folklór. Tanulmányok Verebélyi Kincső tiszteletére*. Budapest: ELTE BTK Folklore Tanszék. 153–160.
- Rafaelli, Sheizaf 1988. Interactivity: From new media to communication. In Hawkins, Robert P. – Wiemann, John M. – Pingree, Suzanne (eds.): *Sage Annual Review of Communication Research: Advancing Communication Science: Merging Mass and Interpersonal Processes*, 16. Beverly Hills: Sage. 110–134.

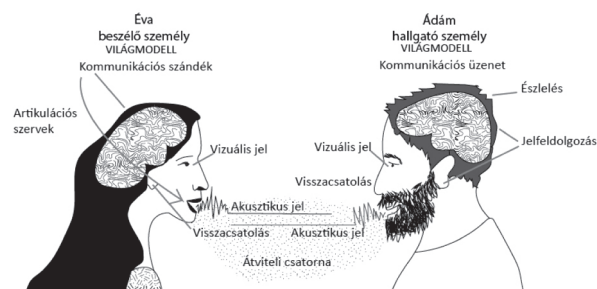
- Ruby, Jay 1973. Up the Zambezi with notebook and camera or being an anthropologist without doing anthropology... with pictures. *Ethnographic Film Newsletter* 4 (3): 12–14. URL: <http://cfaonline.asu.edu/haefer/classes/568/UpTheZambezi.html>. Utolsó letöltés: 2018. március 25.
- Ruby, Jay 1980. Franz Boas and Early Camera Study of Behavior. *Kinesics Report* 3 (1): 7–16.
- Sütheő Péter 1999. *Hypertext. Természetes intelligencia az információ-tudományban*. Budapest: OSZK.
- Szűts Zoltán 2001. *A hypertext*. URL: <http://magyar-irodalom.elte.hu/vita/tszz.html> Utolsó letöltés: 2018. március 25.
- Trigg, Randall 1983. *A Network-Based Approach to Text Handling for the Online Scientific Community*. Ph.D. dissertation. URL: <http://www.workpractice.com/trigg/thesis-chap4.html>. Utolsó letöltés: 2018. március 25.
- Vargha Katalin 2005. Nincs új a net alatt. Antiproverbiumok az interneten. In Gulyás Judit – Tóth Arnold (szerk.): *Mindenek Gyűjtemény II. Tanulmányok Küllös Imola 60. születésnapjára*. Budapest: ELTE BTK Folklore Tanszék. 371–388.
- Vargha Katalin 2016. Miért és hogyan végezzünk online folklorisztikai terepmunkát? In Ispán Ágota – Magyar Zoltán – Mészáros Csaba – Vargyas Gábor (szerk.): *Ethno-Lore. A Magyar Tudományos Akadémia Néprajzi Kutatóintézetének évkönyve. XXXIII*. Budapest: Akadémiai Kiadó. 281–298.
- Wilson, Samuel M – Peterson, Leighton C. 2002. The Anthropology of Online Communities. *Annual Review of Anthropology* (31): 449–467.
- Worth, Sol – Adair, John 1977. Navajo filmesek. In Horányi Özséb (szerk.): *Montázs*. Budapest: Tömegkommunikációs Kutatóközpont. 271–324.

Németh Géza

Kempelentől a WaveNet-ig: a gépi beszédkezelés tudományának fejlődése

1. Bevezetés

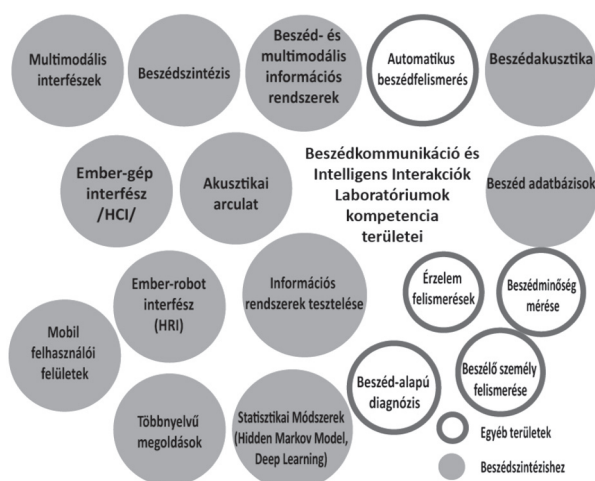
A gépi beszédkezelés a beszédtechnológia tudományterületének egyik ága. Az 1. ábrán láthatjuk a természetes beszédlánc egyszerűsített modelljét. Az emberi kommunikációnak számos alapvető feltétele van. A két partnernek a világról alkotott modellje nagymértékben meg kell egyezzen. Ez a modell hosszú időszaki tanulási folyamata révén alakul ki. A modellhez kapcsolódóan fogalmazódik meg az agyban a beszélő személy kommunikációs szándéka, ami a beszédszerveken keresztül alakul fizikai jelekké (elsősorban akusztikus és vizuális formában). Ezek a fizikai jelek egy átviteli csatornán (természetes közegben a levegőn, gépi megoldásnál hang- vagy videotelefonon) keresztül jutnak el a hallgatóhoz. A hallgató személy érzékszervei adják tovább a megfelelő jelfeldolgozás után az észlelés számára az információt. A kommunikációs üzenet értelmezése a hallgató személy világról



1. ábra. A természetes beszédlánc egyszerűsített modellje

alkotott modelljéhez kapcsolódóan alakul ki. A beszédkommunikáció alapvető jellemzője, hogy a beszélő és a hallgató szerep időről időre felcserélődik, így információelméleti szempontból visszacsatolt rendszerről beszélhetünk. Megjegyzendő, hogy az egészséges beszélő személy saját maga is hallja a beszédét, és ennek is fontos szabályozó szerepe van (például hangerő meghatározásban). A továbbiakban az akusztikus csatorna szerepével foglalkozunk, mert a gépi feldolgozásban általában annak van elsődleges szerepe.

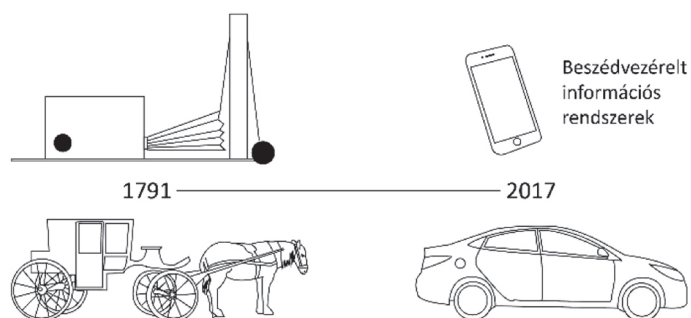
Beszédtechnológiának a természetes beszédlánc egy vagy több elemének gépi megvalósítását tekintjük.¹ A beszédtechnológia interdiszciplináris tudomány, számos bölcsészeti (például nyelvtudomány, fonetika, pszichológia), természettudományi (például fizika, matematika) és műszaki területet (például akusztika, jelfeldolgozás) érint. Laborcsoportunk kompetencia területeit mutatja a 2. ábra.



2. ábra. Laborcsoportunk kompetenciatérületei

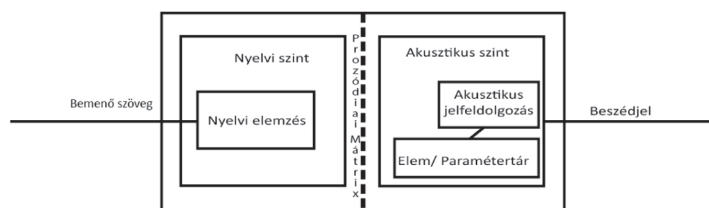
¹ Németh és Olaszky 2010.

A jelen tanulmányban a gépi beszédkeltés tudományának fejlődését tekintjük át. Ezt illusztrálja a 3. ábrán látható idővonal.



3. ábra. A közlekedés eszközeinek és a gépi beszédkeltés technológiáinak fejlődési idővonal

A gépi szövegfelolvasás általánosított modellje a 4. ábrán látható. A nyelvi szinten a bemenetre kerülő szövegből meghatározzuk a kimondandó hangokat és azok alapvető prozódiai jellemzőit (időtartam, intenzitás, zöngés hangokra alapprofrendencia-menet). Az akusztikus szinten pedig a rendelkezésre álló technológiától függő elemtár és jelfeldolgozási módszer segítségével a prozódiai mátrix adatai alapján előállítjuk a kimeneti beszédjelet.



4. ábra. A gépi szövegfelolvasás általánosított modellje

A beszédkeltés gépi modellezése több mint két évszázadra tekinthet vissza.² Hosszú ideig csak az artikulációs csatorna modelljének vezérlését oldották meg. Az 1980-as évek közepéig a megoldások a hangképző szervek (tüdő, légcső, gége, garat, száj- és orrüreg, ajkak) és az artikulációs folyamat működésének leírásán alapultak.³ A hangképzés artikulációs modellezése sikerre vezetett, hiszen a modellel az emberi beszédhez megtevesztésig hasonló hangjelenséget is sikerült létrehozni⁴ a vezérlő paraméterek hosszadalmas kézi optimalizálása révén. A szövegfelolvasáshoz szükség volt a szöveg valós idejű beadására és számítógépes elemzésére is.⁵ Azonban ezzel a megoldással a fő célt, az automatizált gépi szövegfelolvasás emberre emlékeztető szintjét nem sikerült elérni. A géppel keltett beszédjel érthető, de meglehetősen robotos hangzású volt.

Ezért az 1990-es évek elejétől előtérbe kerültek az emberi hangképzés eredményeként előálló hullámforma tárolásán, feldolgozásán, módosításán és visszajátzásán alapuló megoldások.⁶ Ezek segítségével lehetett hosszabb ideig folyamatosan használható gépi felolvasó rendszereket létrehozni (például e-level felolvasása és képernyő felolvasása látássérült emberek számára).⁷ Szűk tématerületen (például időjárás jelentés, menetrend-felolvasás) kutatásaink eredményeképpen már magyar nyelven is lehet az emberi felolvasás minőségét és jellemzőit megközelítő rendszereket létrehozni.⁸ Az elmúlt évtizedben pedig az artikulációs és a hullámforma-alapú megközelítés előnyeinek kombinációját ígérő statisztikai parametrikus beszéd-szintézis (elsősorban Hidden Markov-Model, HMM és Deep Neural Networks, DNN) kialakulásának lehattünk tanúi.⁹ A legújabb technológia

² Kempelen 1989.

³ Stevens, Kasowski és Fant 1953.

⁴ Rosenberg, Schafer és Rabiner 1971.

⁵ Olasz 1989, Klatt és Klatt 1990.

⁶ Moulines és Charpentier 1990, Beutnagel és mtsai 1999.

⁷ Olasz, Németh és Olasz és mtsai 2000.

⁸ Németh, Olasz és Fék 2006.

⁹ Zen, Tokuda és Black 2009, Zen, Senior és Schuster 2013.

(WaveNet)¹⁰ pedig a hullámformából tanulja meg a modell paramétereit, és a megelőző néhány ezer hullámforma minta alapján ad becslést a következő mintára.

Időközben az is kezd körvonalazódni a kutatásokban, hogy az alkalmazási területtől, az ember–gép kapcsolat megoldásától függően változhat a géppel előállított beszéd minőségi követelménye. Például egy beszélő robot esetén az érthetőség a legfontosabb és kimondottan előnyös lehet, ha nem tökéletesen emberi jellegű, hanem robotos hangzású az előállított hang. A robotikából jól ismert a rejtélyes völgy (uncanny valley)¹¹ hatás, mely szerint az emberre hasonlító gép egy bizonyos hasonlósági fokig pozitív érzelmi hatást vált ki, de ezután elérhet egy letörési pontot, ahol már inkább elutasítást vált ki az emberben (zombinak tekintjük). Éppen ezért a tökéletes gépi beszéd létrehozásához és annak elfogadásához nemcsak a beszédkeltés mechanizmusát, hanem az agy működését szemantikai szinten is meg kell(ene) értenünk. Ameddig nem érünk el erre a szintre, addig az éppen aktuális felhasználást figyelembe véve és az a priori rendelkezésre álló információk alapján célszerű a feladathoz illeszteni a gépi beszédkeltés megfelelő változatát. Így lehet optimális ember–gép interfészt megvalósítani.

3. Áttekintés

A gépi beszéd-előállítás tudományos alapjait Kempelen Farkas 1791-ben megjelent könyve fektette le.¹² Az első elektromechanikus beszélőgépet is magyar ember találta fel.¹³ Nagy médianyilvánosságot kapott a Bell Laboratóriumban az 1930-as években fejlesztett elektromechanikus VODER-rendszer.¹⁴ A (nagy)szá-

¹⁰ Zainkó, Tóth és Németh 2017.

¹¹ Mori 1970/2012.

¹² Kempelen 1989.

¹³ Bánó 1916.

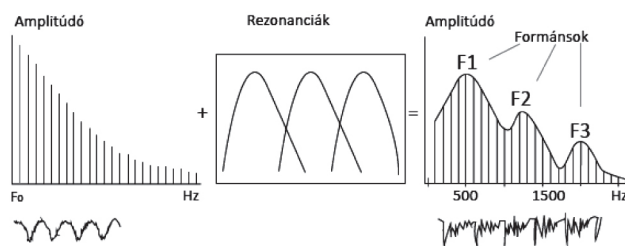
¹⁴ Dudley, Riesz és Watkins 1939.

mítógépes gépi beszédkeltés első megoldásai az 1950-es években születtek.¹⁵ A mini- és mikroszámítógépek megjelenésével a hazai kutatók is követhették a nemzetközi trendeket.¹⁶

Az elmúlt évtizedekben a számítástechnika technológiai fejlődése a gépi szövegfelolvasás területén is több technológiai megközelítés kutatását és alkalmazását tette lehetővé. Látható, hogy a gép beszéd-előállítás témakörében a technológia tükrében mindig változó kutatási kérdések merülnek fel. Ezek megoldása folyamatos kihívást jelent, és egyrészt egymást követő alternatív tudományos generációkat eredményez. Másrészt azzal jellemezhető, hogy a korábbi generációk nem avulnak el (mind a mai napig használatban vannak), hanem az újabb generációk más-más peremfeltételek optimalizálását igénylik és teszik lehetővé.

Formánszintézis

A különböző elvi megközelítések különböző beszédminőséget és gyakorlati alkalmazási lehetőségeket eredményeznek. Az artikulációs¹⁷ megközelítés elsősorban az emberi beszédkeltés mechanizmusainak modellezésére volt alkalmas. A formánsalapú beszéd-szintézissel (lásd 5. ábra) sikerült kötetlen szókészletű, jól



5. ábra. A gépi beszédkeltés formáns modelljének alapelve

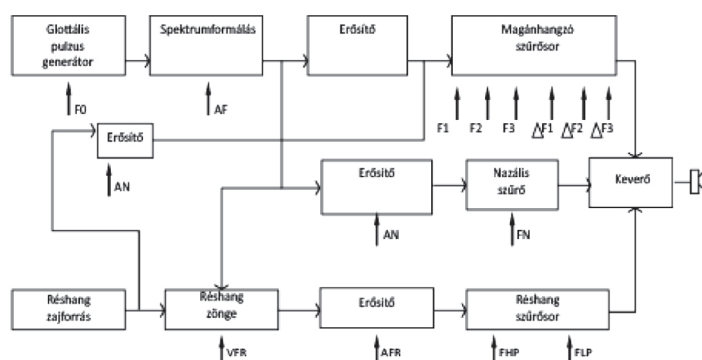
¹⁵ Cooper 1961.

¹⁶ Olasz 1978, Gordos és Takács 1983.

¹⁷ Mermelstein 1973.

érthető, kereskedelmi forgalmazásra alkalmas, de egyértelműen gépies hangzású, gépi hangot előállítani.

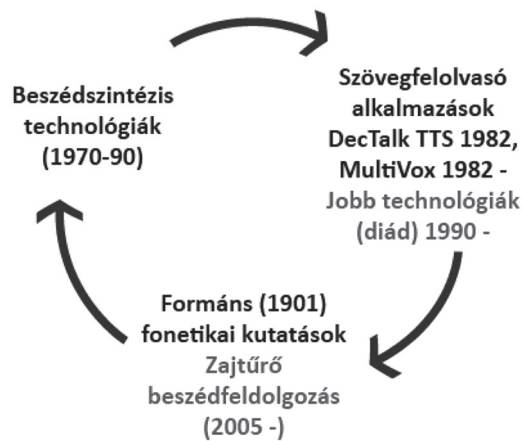
A modell lényege az ún. forrásszűrő megközelítés. A zöngés hangokat azonos alapprofrendiciájú (F_0) periodikus gerjesztéssel, a zöngétleneket fehérzaj-szerű forrásjellel és az artikulációs csatornát szimuláló szűrősorral modellezzük. Az így kapott kimeneti jel hullámformája és frekvencia spektruma (azokban a formáns értékek, melyek meghatározóak a magánhangzók észlelésében) jó közelítéssel megegyezik a természetes beszéddel. A 6. ábrán egy formánsmodell részletes blokkdiagramját láthatjuk.



6. ábra. Formánszintetizátor blokkdiagramja

A mai napon (2018. március 14.) bekövetkezett haláláig ilyen elveken alapuló rendszert használt Stephen Hawking, az ismert elméleti fizikus, mert az évtizedek alatt azonosult a gép hangkarakterével.

Érdeemes röviden áttekinteni azt a ciklikusságot (7. ábra), ami az elméleti kutatások alapján elért technológiai eredmények után a gyakorlati alkalmazásokig jut, majd az itt felmerülő problémák újabb elméleti megalapozottságot igényelnek.



7. ábra. A formáns témakör ciklikus fejlődése

A formánsokkal kapcsolatos elméleti kutatások a 19. század végétől erősödtek fel,¹⁸ és az 1970-es évektől eredményeztek gyakorlati alkalmazást is ígérő beszédszintézis technológiákat.¹⁹ Talán a legismertebb formáns alapokon nyugvó angol nyelvű szintetizátor a DecTalk volt, ami az MIT professzora, Dennis Klatt kutatásain alapult.²⁰ Stephen Hawking is ennek egy változatát használta.

Érdemes megjegyezni, hogy egy ilyen önálló dobozban található eszköz ára 1984-es megjelenésekor mintegy 4000 USD volt. Ezzel gyakorlatilag egy időben készült el a HungaroVox rendszer az MTA Nyelvtudományi Intézetében²¹ (lásd 8. ábra).

Ennek továbbfejlesztéseként készültek el a MultiVox különböző változatai²² a BME-n előbb PC-hez illesztett hardver, illet-

¹⁸ Lloyd 1890.

¹⁹ Flanagan és Rabiner (eds.) 1973.

²⁰ D. Klatt 1987.

²¹ Kiss és Olasz 1984.

²² Olasz 1989.



8. ábra. A HungaroVox rendszer doboza

ve önálló dobozos eszközként, majd szabadon letölthető tisztán szoftver változatban.

A formánsokat nemcsak a beszéd szintézisére, hanem felismerésére is alkalmazták.²³ A formánsok automatikus meghatározása és transzformációja is a folyamatosan vizsgált elvi problémák közé tartozik.²⁴ Újabban sokféle szintetizált beszédhang előállításában és a zajos beszéd felismerésénél erősödött fel az érdeklődés ebben az irányban.



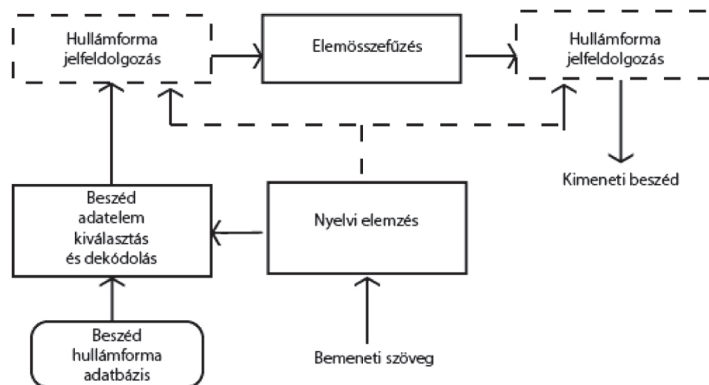
9. ábra. A MultiVox rendszer doboza

²³ Furui 2005.

²⁴ Bóhm és Németh 2006.

Elemösszefűzéses és elemkiválasztásos gépi szövegfelolvasás

A korábbi parametrikus (elsősorban formánsalapú) érthető, de erősen robotos hangzású szövegfelolvasási technológia továbbfejlesztésére a 80-as évek végétől alakult ki az a koncepció, hogy próbálkozzunk természetes beszéd rögzítésével, címkézésével és a visszajátszáskor a megfelelően kiválasztott elemek összefűzésével, és (ha szükséges) jelfeldolgozás segítségével történő optimalizálásával. Ennek egyik lehetséges megoldását láthatjuk a 10. ábrán. Ennek a megoldásnak az egyik formája, hogy alapelemnek hangpárokat reprezentáló beszédhullámforma-részleteket (ún. diádok, angolul diphone) választunk. Ekkor például az alma szót _a, al, lm, ma, a_ (_ a szünet jele) diádokból lehet előállítani. A magyar nyelv 39 hanggal (25 mássalhangzó –C- és 14 magánhangzó –V-) plusz a szünet (_ jel) lefedhető, tehát az adatbázisban mintegy $40^2 = 1600$ elemre van szükség. A hosszú mássalhangzókat időtartam módosítással tudjuk megoldani. A minőséget egyrészt a sok, folytonossági hibát okozó vágási pont, másrészt a prozódia megvalósító modell egyszerűsége és a jelfel-



10. ábra. Elemösszefűzéses és elemkiválasztásos gépi szövegfelolvasó rendszer általános blokkdiagramja (a – blokkok opcionálisak)

dolgozás korlátozza.²⁵ A minőséget tovább javíthatja a hangok vágás nélküli összefűzését legalább a magánhangzók esetében figyelembe vevő hanghármasok (triád, angolul triphone) alkalmazása. Például ekkor az alma szót az _al, lm, ma_ két triádból és egy diádból lehet előállítani.

Ezzel a megoldással egyrészt az eredeti emberi hangszínre emlékeztető gépi beszédet lehet létrehozni, másrészt viszonylag kis számú számítási kapacitás mellett lehet változtatható hangkaraktereket (változó alaphangfrekvencia és beszédtempó) kialakítani. Ennek különös fontossága van a látássérült emberek kommunikációjának szempontjából.

Az elemösszefűzéses megoldás hátránya, hogy egy-egy hangkapcsolathoz csak egyetlen emberi bemondásból származó mintát tárol. Ezért a beszéd megfelelő prozódiai jellemzőit az összefűzött elemeken végzett jelfeldolgozással kell biztosítani.

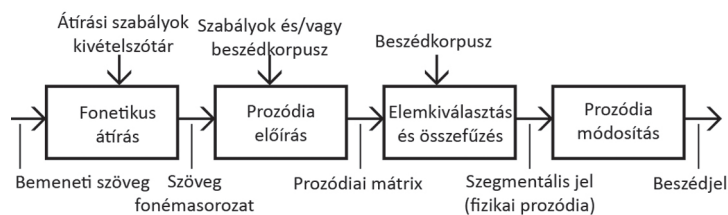
Elemkiválasztásos gépi szövegfelolvasás

Az első magyar nyelvű korpuszalapú, hullámforma elemválogatásra épülő gépi szövegfelolvasó rendszer modelljét láthatjuk a 11. ábrán. Szubjektív értékelés szerint az emberi felolvasáshoz jobban hasonlító hangot állít elő, mint a korábbi magyar nyelvű szövegfelolvasó rendszerek. Három célorientált alkalmazási területen (időjárás-jelentés, árlista és pályaudvari tájékoztató felolvasása) kiderült, hogy ennek a technológiának a felhasználásával lehetséges az emberi felolvasáshoz meglehetősen hasonló magyar nyelvű gépi felolvasást létrehozni.

A 90-es évek második felében kezdett megfogalmazódni az a koncepció, amit korpuszalapú beszéd-szintézisnek nevezünk.²⁶ Az elképzelés alapötletét az az általánosan elfogadott elv adja, hogy egy hullámforma-összefűzésen alapuló szövegfelolvasó rendszer minőségét döntően az összefűzések száma és az összefűzött ele-

²⁵ Olasz, Németh és Olasz és mtsai 2000.

²⁶ Möbius 2000.



11. ábra. Korpusz alapú, hullámforma elemkiválasztásos beszédszintetizátor modellje (Fék és mtsai 2006)

meknek az emberi kiejtéshez való hasonlósága határozza meg. Minél hosszabb elemekből állítjuk elő a szintetizált beszédet – az összefűzési pontok számának csökkenése és a természetes beszédhez való jobb illeszkedés miatt –, annál jobb lesz az elért minőség. Az ideális tehát az lenne, ha minden lehetséges felolvasandó szöveg, de legalábbis minden lehetséges mondat szerepelne elemként a rendszer adatbázisában. Természetesen ez a gyakorlatban kivitelezhetetlen, ezért ennél rövidebb egységeket vesznek fel az adatbázisba, de azzal a céllal, hogy nagy valószínűséggel hosszú elemekből összefűzhető legyen a kimenet.

Ennek egyik szélsőséges megoldása például az autóbuszokon alkalmazott bemondások digitális rögzítése, majd megfelelő egyszerű vezérlés (például nyomógombok) segítségével történő visszajátszása. Például: *A következő megálló – a Keleti pályaudvar.* A mondat első fele a hangos bemondásban a rögzített elemet képviseli, a mondat második eleme a változót. Fontos látni, hogy az ilyen összeillesztéseknél a prozódianak illeszkednie kell egymáshoz. Ez a példában azt jelenti, hogy a rögzített rész mindig az üzenet kezdete, a megálló neve pedig a vége (ha megcserélnénk a kettőt, és úgy játszanánk le, akkor prozódiaileg természetellenes hangzást kapnánk). Természetesen ennek a megoldásnak egyrészt jelentős a tárgyigénye, másrészt erősen korlátozott a témaköre.

A fenti koncepció alapján külföldön már készült néhány korpuszalapú beszédszintetizátor a világnyelvekre,²⁷ magyar megol-

²⁷ Möbius 2000, Kawai és mtsai 2004.

dást azonban elsőként a BME-n hoztunk létre. Munkánk során felhasználtuk a korábbi magyar nyelvű kutatások²⁸ eredményeit is. Kutatásaink során arra a fő kérdésre kerestük a választ, hogy lehetséges-e olyan gépi beszédkeltési modellt létrehozni magyar nyelvre, ami akár az emberi bemondásra megtévesztésig hasonló kimenetet tud létrehozni kötött, de nagy változatosságot tartalmazó tématerületen. A más nyelvekre kidolgozott modellek nem feltétlenül hasznosíthatók, hiszen a magyar nyelv ragozó jellege miatt például az angol nyelvre kidolgozott szóalapú megközelítések nem alkalmazhatók közvetlenül.

Első kísérleti területünk az időjárás jelentés témaköre volt. Húsz internetes oldal 2004 áprilisa és 2005 májusa közötti időjárás-jelentéseinek alapján reprezentatív szöveges adatbázis jött létre (56 000 mondat, 670 000 szó szintű szövegelem). Ez a szöveges adatbázis túl nagy ahhoz, hogy reális erőforrások mellett (legfeljebb néhány hét alatt) egy professzionális bemondó felolvassa. A méret csökkentésére a következő modell vált be: ne csak az előforduló mintegy 5200 szóalak és a számok jó minőségű felolvasásához szükséges mintegy 230 számelem egy-egy változata kerüljön be a szűkített szöveges adatbázisba, hanem a későbbiekben részletezett prozódiai változatosság is megoldott legyen.

A szintézis optimális alapelemének a szóelemet választottuk (két szóköz közötti karaktersorozat), valamint az ebből felépülő hosszabb szövegrészeket (szófüzér, önálló mondatrész stb.). A szó méretű elem egyrésztől hosszabb a diád-triád elemeknél, tehát akusztikai tartalma biztosan jobban képviseli az optimális hullámformát, másrésztől a percepció feldolgozásunk során az anyanyelvi bázisunk a szó feldolgozására épül az agyunkban. Ha tehát jó akusztikai tartalmú szó kerül a szintetizálendő mondatba, akkor természetesebb hangzásúnak fogjuk ítélni, mint a diád/triádokból összerakott ugyanazon szót. Mindezt segíti, ha prozódiai szempontból is megfelelő szó kerül a szintetizálendő mondat adott helyére. Mindezekből adódik, hogy a szintézishez

²⁸ Olaszky és Németh 1999.

használt beszédadatbázisnak két kritériumnak kell eleget tennie. Az első az, hogy minden szóból legalább háromfélét kell tartalmaznia (mondatkezdő, -belső és -záró elem). A második az, hogy tartalmazzon megfelelő diád/triád lefedettséget is tetszőleges szöveg (az optimálisnál rosszabb, de lehetséges) előállításához

A prozódia modellben az alapegység a mondat. A modell szorosan összefügg a szintetizálendő szöveg szerkezetével. Jelen esetben döntően kijelentő mondatokat modellezünk. A kijelentő mondat prozódiai szerkezete jól körülhatárolható, ismert egységekből áll. Ezeket az egységeket a mondaton belüli hely szerinti pozicionálással (hol van a szó a mondatban), valamint a központoszással (vesszők, gondolatjelek stb.) kapcsolatba lehet hozni. Ez a modell lényege. Ugyanaz a modell kerül alkalmazásra a szöveges adatbázisban, a beszédadatbázisban és a szintetizálendő mondatban is. Alkalmazásával nincs szükség prozódiai jellegű jelfeldolgozás használatára a szintézis során. Az ezeknek a peremfeltételeknek megfelelően mohó algoritmussal²⁹ kialakított szöveges adatbázis rugalmasan bővíthető. Végül 5821 mondatot, 102 940 szót tartalmaz, ami 488 093 hangnak (fonémának) felel meg.

A szöveges adatbázist mintegy két hónapos munkával egy professzionális bemondó felolvasásában rögzítettük. Ezután utófeldolgozás következett. A hullámformát több szinten címkéztük. A legelső szinten fonéma (hang) címkékkel történő eljáráshoz félautomatikus eljárás valósult meg a BME TMIT-en fejlesztett beszéd felismerő³⁰ felhasználásával. A beszéd felismerőt ún. kényszerített üzemmódban (forced alignment, az ismert szövegnek megfelelő hangok pozícióját kellett megjelölni a hullámformában) használtuk.

A prozódiai modulban a magyar nyelvhez és a célorientált megközelítéshez illeszkedő új indirekt eljárást valósult meg. Az adott szó mondatban, illetve prozódiai egységben elfoglalt helyéhez lehetőleg optimálisan illeszkedő elemek (elsősorban szó,

²⁹ Cormen, Leiserson és Rivest 1990.

³⁰ Mihajlik és mtsai 2007.

ha az nincs, diád/triád hangelemek) összefűzését végzi az algoritmus.

Az elemkiválasztás és összefűzés modulban két költségfüggvény összegének minimalizálása valósul meg új, fonetikai szempontok szerint kialakított költségfüggvények alapján. Az egyik költségfüggvény az egyes elemek (szó- és hangszinten eltérő) egymáshoz illeszkedésének (folytonosságának) felel meg (ún. összefűzési költség). Mivel a kiejtés folyamatos, a (szó)határon törekedni kell arra, hogy a spektrális illeszkedés (például formánsmenet) is folyamatos legyen. A szavak első és utolsó hangjának illeszkedése kerül vizsgálatra, és az illeszkedés költsége több szempont alapján számítható ki. Magas költségű például, ha a szóhatáron magánhangzók találkoznak (dunántúli áramlások). Az ilyen szavak magas költséget képviselnek. Nulla a költség, ha a két szó egymás mellett helyezkedik el a beszédkorpuszban, hiszen ekkor a csatlakozásuk is optimális. Ebből adódik, hogy akkor nagyon optimális a keresés, ha nem szavakat, hanem szófüzereket találunk a korpuszban. Az esetek nagy részében (ha a beszédkorpusz elég nagy) ez meg is valósul, így a szintetizált szöveg hangzása közel lesz a természeteshez.

A másik költségfüggvény határozza meg, hogy hangsor és hangkörnyezet szempontjából a kiválasztott elem (szó, szófüzér vagy hang) mennyire felel meg a prozódiai követelményeknek. Itt szempont az is, hogy a kiválasztott elem a mondatkorpusz ugyanazon mondatában szerepel-e, mint az előző. Ha igen, akkor a költséget ez a tény is csökkenti. A prozódiai költség meghatározásánál – az időtengelyi pozíción felül – felhasználjuk az alapfrekvencia (F_0) értékének a változását is. Ha nagy F_0 ugrás van a két elem között, akkor a költség magas lesz, tehát a két elem nem illeszthető össze.

A költségfüggvények súly értékeit iteratív módon, mintegy 500 mondat többszöri szintézisével határoztuk meg. A költségfüggvények alapján először a szószintű, majd a hangszintű optimális elemeket választjuk ki Viterbi-algoritmus segítségével. Ha a költségfüggvény-optimalizálás ellenére csak jelentős illesztetlenséget tartalmazó elemeket találunk a felolvasandó szöveghez,

akkor kerül sor a prozódia simítását végző modul alkalmazására. Ez mindenképpen jeltorzulást okoz, és gyakran jól hallható a kimeneten.

Ideális esetben prozódiai módosítást végző jelfeldolgozás nélkül történik az összefűzés.³¹ Ebben a megoldásban egy-egy hangkapcsolathoz akár több tízezer minta is tartozhat. A nehézsége a megnövekedett tárhelyigényen túlmenően az, hogy a hatalmas számú alternatíva közül kell valós időben megtalálni a (közel) optimális megoldást.

A hullámforma alapú megoldások hátránya, hogy a legjobb minőséghez minden egyes beszélőtől külön-külön beszédadatbázist kell rögzíteni. Beszédadatbázison a következőt értjük: hanganyag, azaz a felolvasásból származó emberi beszéd, az elhangzott szöveg fonetikus átírata és többszintű szegmentálási címkék halmaza. A beszédadatbázist (más néven beszédkorpuszt) jellemzően az adott kutatási feladathoz illesztve készítik el. Ez nemcsak hangfelvételt jelent, hanem az adott technológiától függő összetett címkézési feladatot is. Ha a beszédtempót változtatni akarjuk, az csak jelfeldolgozási műveletekkel lehetséges, amik rontják a beszédjel minőségét. Ennek a technológiának a legújabb hazai felhasználási területe a MÁV-állomásokon hallható magyar és angol nyelvű célorientált gépi szövegfelolvasó rendszer.³² A mai napig ez a technológia szolgáltatja a legjobb beszédminőséget.

Statisztikus parametrikus gépi felolvasó rendszerek

A gépi beszédkeltés terén az elmúlt években – számos előnyének köszönhetően – a statisztikai parametrikus beszéd-szintézis vált az egyik legaktívabb kutatási területté.³³ Ennek során először ki-nyerjük a jellemző paramétereket (például spektrális összetevők,

³¹ Möbius 2000, Németh, Olasz és Fék 2006.

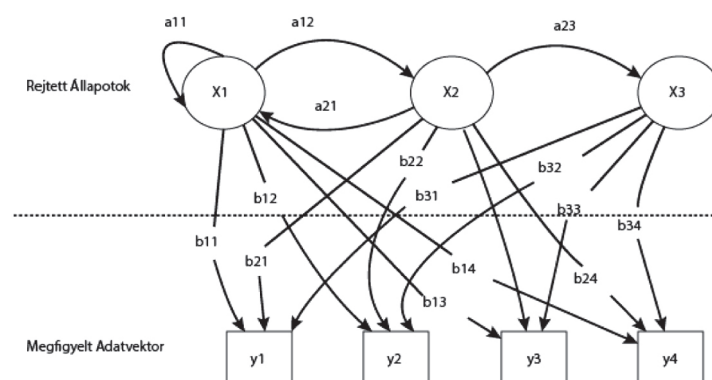
³² Zaikó, Bartalis és mtsai 2015.

³³ Zen, Tokuda és Black 2009.

alapfrekvencia, hangidőtartamok, hangok elhelyezkedése, hangkörnyezet) a beszédkorpuszból, majd ezen paraméterek sokaságát HMM és DNN modellekkel helyettesítjük.

Rejtett Markov-modell alapú
gépi beszédeltés

Jellemzően a beszéd felismerésben már több évtizede sikeresen alkalmazott rejtett Markov-modell (HMM), valamint az újabban előtérbe került Deep Neural Networks (DNN) alapú megközelítés a legelterjedtebb ebben a modellalkotásban. A 12. ábra szemlélteti a HMM-modell alapgondolatát.

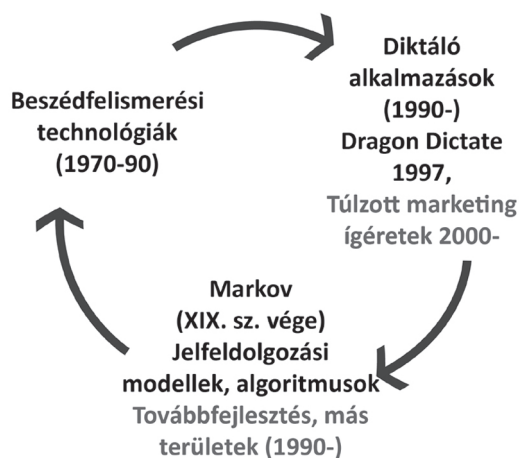


12. ábra. A HMM-modell alapgondolata

Az y_1, y_2, \dots adatvektorokat tudjuk megfigyelni. Ezeket az x_1, x_2, \dots állapotok közti átmenetek során emittálja a modell. Az állapotokat nem ismerjük, ezért rejtett a modell. Az állapot átmenetek valószínűségét adják meg az $a_{11}, a_{12}, a_{21}, \dots$ súlyok. A $b_{11}, b_{12}, b_{13}, b_{14}$ valószínűségek azt jelzik, hogy az adott állapotban milyen valószínűséggel bocsátja ki a modell a megfelelő adatvektort.

Tehát a feladat az, hogy az adatvektorok ismeretében becsüljük meg, hogy milyen állapotátmenet-sorozat valósult meg a modellben. Beszédfelismerés esetén valamilyen lényegkiemelt paraméter (például cepstrális együtthatók) az adatvektorok, az állapotok pedig a kimondott beszédhangoknak felelnek meg. Beszédszintézis során pedig az adatvektorokat a bemeneti szöveghez tartozó intonációs mátrix adatai jelentik, az állapotoknak pedig egy beszédkódoló paramétervektorai (egyebek között spektrális adatok) felelnek meg.

A HMM témakörben is megfigyelhető az alapkutatás – technológia kutatás-fejlesztés – alkalmazások ciklikussága (13. ábra). Az alapelveket Andrej Markov, orosz matematikus dolgozta ki a 19. század végén és a 20. század elején.³⁴ Az 1960-as években merült fel az elmélet gyakorlati felhasználása.³⁵ Az IBM-nél Fred Jelinek és kutatócsoportja dolgozta ki



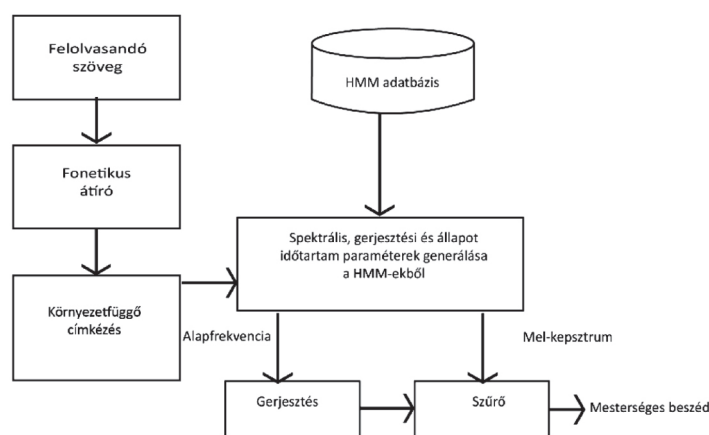
13. ábra. A HMM-kutatások ciklusai

³⁴ Markov 1913.

³⁵ Baum és Petrie 1966.

ezen elmélet alapján az első gépi beszéd felismerő rendszert a 70-es években.³⁶ Ennek alapján jöttek létre az első kereskedelemben kapható nagyszótárú beszéd felismerő rendszerek (IBM Tangora, Dragon Systems, Philips dictation stb.). A beszéd felismerésben elért sikerek vezettek oda, hogy felmerült az elmélet alkalmazása gépi szövegfelolvasás céljaira is. Az első ilyen rendszert a nagoyai egyetemen Tokuda professzor irányításával fejlesztették ki.³⁷ Természetesen ez jelentős további alap kutatási feladatokat vetett fel.

Egy HMM-alapú gépi szövegfelolvasó rendszer blokkdiagramját látjuk a 14. ábrán.³⁸ A rendszer előnye, hogy megfelelően felcímkezett adatbázisból automatikusan állítható elő a HMM-modellek adatbázisa. Ez jelentős számítási időt igényel. Viszont a modellek már gyorsan generálják a beszéd kódoló vezérléséhez szükséges paramétersorozatot. Ez lényegesen hatékonyabb, mint

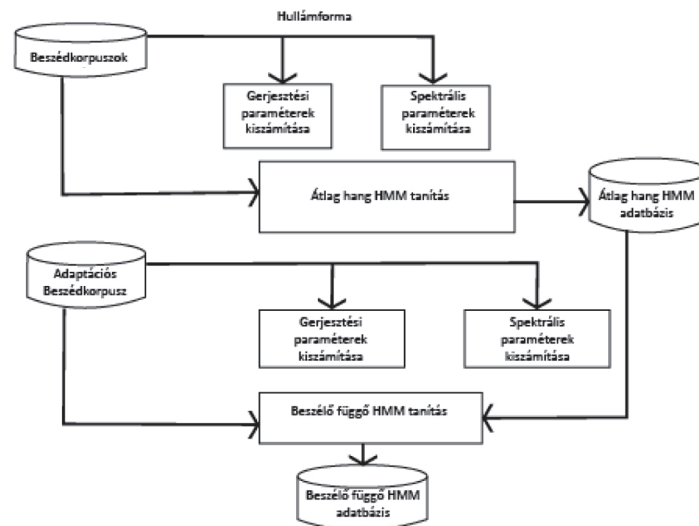


14. ábra. A HMM-szintézis alapelve

³⁶ Jelinek 1976.

³⁷ Tokuda és mtsai 2000.

³⁸ Tóth és Németh 2008.



15. ábra. A HMM-hangadaptáció blokkdiagramja

a hullámforma elem összefűzéses vagy elemkiválasztásos megoldások. Így megnyílt az út a sok hangon beszélni képes gépi szövegfelolvasó rendszerek kialakítása előtt.

További előnyöket jelent az, hogy viszonylag rövid (akár néhány percnyi) hangminta alapján is lehetséges az adott beszélő hangjára emlékeztető felolvasó rendszer létrehozása.³⁹ Ehhez célszerűen mintegy tíz beszélő személy nagyobb (személyenként kb. 2000 mondatot tartalmazó) hangadatbázisából egy ún. átlag hang-HMM-modellt hozunk létre az adott nyelven. Majd ezt a modellt tudjuk a rövid felvétel alapján az adott beszélőhöz igazítani. Ennek a megoldásnak a blokkdiagramja látható a 15. ábrán.

³⁹ Tóth és Németh 2010.

Neurális hálózatok

A neurális hálózatok elmélete is meglehetősen hosszú időre tekinthet vissza.⁴⁰ A gyakorlati alkalmazáshoz nagy lökést adott a modellek tanítására kidolgozott eljárás.⁴¹ A számítástechnika gyors fejlődése is népszerűsítette ezt a megközelítést a 80-as években. A beszédfelismerés területén a HMM versenytársaként tekintettek erre az alternatívára.⁴² Azonban az adott technikai korlátok között a HMM jobb eredményeket adott. Ezért erősen lecsökkent a téma iránti lelkesedés, de az elméleti munka tovább folytatódott. Többek között ez is vezetett az IBM DeepBlue rendszere kifejlesztéséhez, ami 1997-ben legyőzte az aktuális sakkvilágbajnokot.⁴³ A számítási kapacitás növekedése (különösen a grafikus jelfeldolgozó kártyák – GPU – megjelenése) és a memóriaméreték megsokszorozódása elhozta a sok réteget tartalmazó mély neurális hálók (deep neural networks, DNN) felhasználási lehetőségét is a nagyszótárú gépi beszédfelismerés területén a HMM-nél jobb eredményekkel.⁴⁴

Ezután természetesen kezdtek el alkalmazni a DNN-megoldásokat a gépi szövegfelolvasás területén is, például a spektrális és a prozódiai paraméterek becsléséhez.⁴⁵ Ezt a folyamatot illusztrálja a 16. ábra.

A memória és GPU-kapacitás rohamos növekedése és a felhőalapú infrastruktúrákban összpontosuló hatalmas erőforrások vezettek korábban elképzelhetetlen számítás igényű modell tanítási módszerek kidolgozásához. 2016 januárjában jelent meg az első publikáció arról, hogy pixelenként tanítanak be pixelbecslő neurális hálózatot.⁴⁶ Ugyanezen év szeptemberében ez az elv már

⁴⁰ McCulloch és Pitts 1943.

⁴¹ Werbos 1974.

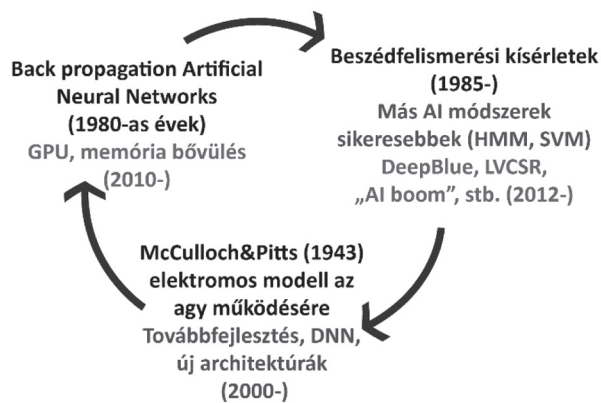
⁴² Waibel és mtsai 1989.

⁴³ Campbell, Hoane és Hsu 2002.

⁴⁴ Dahl és mtsai 2012.

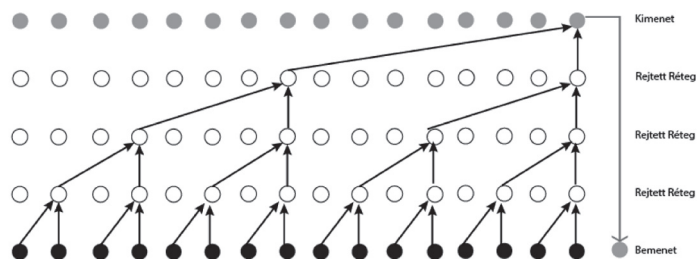
⁴⁵ Zen, Senior és Schuster 2013, Nagy és Németh 2016.

⁴⁶ Van den Oord, Kalchbrenner és Kavukcuoglu 2016.



16. ábra. A neurális hálózati kutatások ciklusai

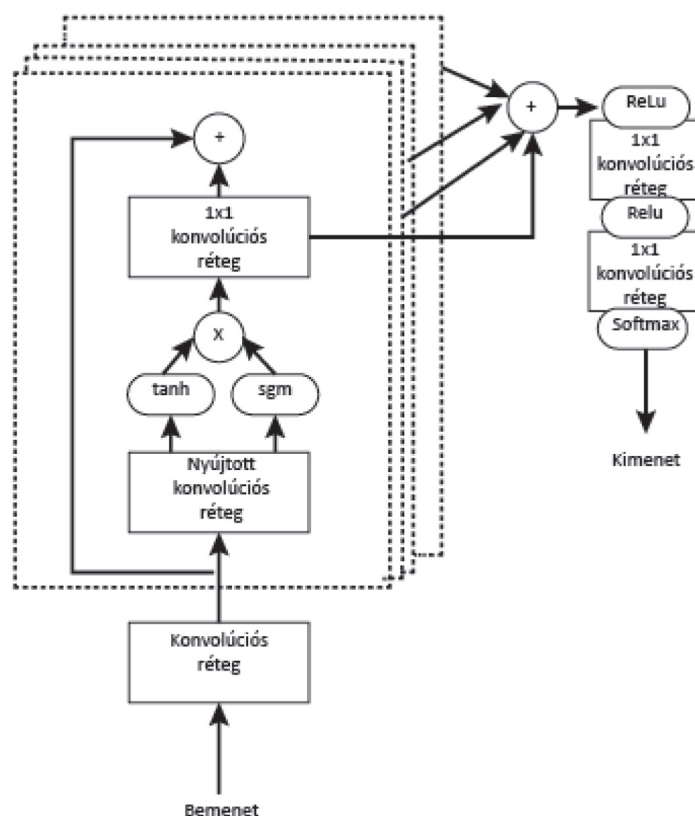
a gépi szövegfelolvasás témakörében került alkalmazásra.⁴⁷ Nem lényegkiemelt paramétereket, hanem a következő mintát becslő néhány ezer előző minta alapján a hálózat (lásd 17. ábra).



17. ábra. A WaveNet alapelve
(az ábra forrása: <https://deepmind.com/blog/wavenet-generative-model-raw-audio/>)

A beszéd mellett zene generálására is alkalmazható. Büszkék vagyunk arra, hogy kutatócsoportunk néhány hónap alatt magyar

⁴⁷ Van den Oord és mtsai 2016.



18. ábra. Magyar WaveNet TTS blokkdiagram
(Zainkó, Tóth és Németh 2017)

nyelvre is alkalmazta ezt a megközelítést⁴⁸ a 18. ábrán látható felépítésben.

A terület rohamos fejlődését jelzi, hogy a Google személyi asszisztensében 2017 szeptemberében (az első cikk megjelenése után nagyjából egy évvel) az amerikai angol és a japán nyelvű termékek ezen megközelítés alapján generált hangon szólalt meg.

⁴⁸ Zainkó, Tóth és Németh 2017.

3. Meghallgatásos tesztek, értékelések

A gépi szövegfelolvasás és a felhasználói felületek értékelésében általánosan elterjedt az eredmények MOS (Mean Opinion Score) és CMOS (Comparison Mean Opinion Score) alapú értékelése. MOS-alapú teszt esetén a tesztalanyok a mintákat 1-től (legrosszabb) 5-ig (legjobb) értékelhetik (egész számokkal), CMOS esetén pedig szintén ötelemű skálán két minta közül kell a tesztalanyoknak eldönteniük, hogy melyik minta tesz jobban eleget a teszt osztályozási kritériumának (például minőség, természetesség, érthetőség). A tesztek során bizonyos esetekben a minőség fogalom értelmezését a tesztalanyokra bízjuk. Ekkor az osztályzás általános visszajelzést ad arról, hogy a tesztalanyok mennyire tartják jónak vagy rossznak az adott rendszert. Ez esetben a rendszer értékelésében számos paraméter, például természetesség, érthetőség, a hang által tesztalanyban keltett érzélem stb. szerepet játszik.

A tanulmányban ismertetett technológiák hangmintái meghallgathatók a <http://smartlab.tmit.bme.hu> és a <http://magyarbeszed.tmit.bme.hu> honlapokon.

4. Az eredmények alkalmazhatósága

Az áttekintett gépi beszédkeletési technológiák ilyen vagy olyan szempontból ma is használhatók, de újabb elméleti és gyakorlati problémákat vetnek fel. Tehát szó sincs arról, hogy a beszéd-szintézis témaköre megoldottnak lenne tekinthető. Néhány megoldásra váró kutatási probléma: az emberi beszédhez hasonló változatosság (minden emberi megszólalás egyedi és megismételhetetlen), az adott kommunikációs kontextushoz illő beszédstílus alkalmazása, gyors adaptáció új témakörökhöz.

Viszont már évtizedek óta vannak érdemi gyakorlati eredmények. A cikkben bemutatott BME kutatási eredmények többek között a következő területeken kerültek alkalmazásra:

- a www.metnet.hu időjárásportál, illetve a Microsoft 2013-as fejlesztői versenyén nyertes Időjárás Mindenkinek Windows8 alkalmazás,
- számos MÁV-állomás hangos utastájékoztató rendszere,
- egy távközlési szolgáltató árlista bemondó szolgáltatása,
- egy távközlési szolgáltató automatizáltan kialakított interaktív hangválasz (IVR) rendszere,
- beszéd-dialógus mintarendszer intelligens lakás prototípusban a BelAmi projekt keretében,
- VoxAid2006 prototípus siketnéma emberek telefonálásának támogatására,
- VoxAid2012 prototípus beszédsérült emberek mindennapi kommunikációjának támogatására, logopédiai és afáziás betegek rehabilitációjának támogatására,
- prototípusrendszer mobil és autós információs szolgáltatásokhoz Android platformon,
- beszédvezérelt okostévé-készülék prototípus,
- beszélő mobil alkalmazások vak emberek számára Symbian, Windows Phone és Android platformon.

Köszönetnyilvánítás

Köszönöm elsősorban a BME TMIT Beszédkommunikáció és Intelligens Interakciók Laborcsoport csapatmunkáját, másrészt a BME TMIT munkatársainak, hallgatóimnak és kutatási partnereinknek az együttműködését. Az ábrák formázásában Németh Zsuzsanna volt segítségemre.

A cikkben áttekintett hazai kutatások eredménye többek között a BelAmi, TÁMOP-4.2.1/B-09/1/KMR-2010-0002, CESAR (ICT PSP No 271022, EU_BONUS_12-1-2012-0005), PAELIFE (AAL_08-1-2011-0001), VUK (AAL-2014-1-183), DANSPLAT (Eureka 9944) valamint az EITKIC_12-1-2012-0001 projekt keretében jöttek létre (a projektek a Kutatási és Technológiai Innovációs Alap támogatásával valósultak meg).

Irodalom

- Bánó Miklós. Tetszőleges szöveg reprodukálására alkalmas beszélőgép. Magyarország Szabadalom száma: 74361 . 1916. 11 30.
- Baum, Leonard E – Ted Petrie 1966. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.* 1554-1563.
- Beutnagel, Mark – Conkie, Alistair – Jürgen, Schroeter – Stylianou, Yannis – Syrdal, Ann 1999. The AT&T next-gen TTS system. In Joint Meeting of ASA, EAA, and DAGA. 18–24.
- Bőhm Tamás – Németh Géza 2006. Algoritmus formánsok követésére, módosítására és szintézisére. *Híradástechnika* LXI (8): 11–16.
- Campbell, Murray – Hoane, A. Joseph Jr – Hsu, Feng-hsiung 2002. Deep Blue. *Artificial Intelligence* 57–83.
- Cooper, Franklin S. 1961. Speech synthesizers. The Hague: Mouton & Co. Proceedings of the 4th International Congress of Phonetic Sciences, (Helsinki), 1961.
- Cormen, Thomas H. – Leiserson, Charles E.– Rivest, Ronald L. 1990. Chapter 17 Greedy Algorithms. In *Introduction to Algorithms*, 768. Mcgraw-Hill.
- Dahl, George E. – Dong, Yu – Deng, Li – Acero, Alex 2012. Context-Dependent Pre-Trained Deep Neural. *IEEE Transactions on Audio, Speech, and Language Processing* 20 (1): 30-42.
- Dudley, Homer. – Riesz, R. R. – Watkins, S. A. 1939. A Synthetic Speaker. *J. Franklin Inst.* 227. 739–764. (Reprinted in Flanagan and Rabiner 1973).
- Fék Márk – Pesti Péter – Németh Géza – Zainkó Csaba. Generációváltás a beszéd szintézisben. *Híradástechnika* LXI (3): 21–30.
- Flanagan, James – Rabiner, Lawrence (eds.) 1973. *Speech Synthesis*. Pennsylvania: Dowden, Hutchinson & Ross, Inc.
- Furui, Sadaoki 2005. 50 Years of Progress in Speech and Speaker Recognition Research. *Ecti Transactions on Computer and Information Technology*. 64–74.
- Gordos Géza – Takács György 1983. *Digitális beszédfeldolgozás*. Budapest: Műszaki Könyvkiadó.
- Jelinek, Frederick 1976. Continuous speech recognition by statistical methods. *Proc. IEEE* 64: 532–536.

- Kawai, Hisashi – Toda, Tomoki – Ni, Jinfu – Tsuzaki, Minoru – Keichi, Tokuda 2004. Ximera: a new TTS from ATR based on corpus-based technologies. *Proc. of the 5th ISCA Speech Synthesis Workshop*. Pittsburgh. 642–645.
- Kempelen Farkas 1989. *Az emberi beszéd mechanizmusa, valamint a szerző beszélőgépezének leírása*. Budapest: Szépirodalmi Könyvkiadó.
- Kiss Gábor – Olasz Gábor 1984. A Hungarovox magyar nyelvű, szótár nélküli, valós idejű párbeszédész beszéd szintetizáló rendszer. *Információ Elektronika* 19 (2): 98–111.
- Klatt, Dennis H – Klatt, Laura C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87 (2): 820–857.
- Klatt, Dennis 1987. How Klattalk became DECTalk: An Academic's Experiences in the Business World. *Proc. of Speech Tech '87*. New York: Media Dimensions Inc. 293–294.
- Lloyd, Richard J. 1890. *Some Researches into the Nature of the Vowel-Sound*. Liverpool: Turner and Dunnett.
- Markov, Andrey A. 1913. An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains. *Bulletin of the Imperial Academy of Sciences of St. Petersburg*. 153–162.
- McCulloch, Warren – Pitts, Walter H. 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* 5: 115–133.
- Mermelstein, Paul 1973. Articulatory model for the study of speech production. *Journal of the Acoustical Society of America* 53 (4): 1070–1082.
- Mihajlik, Péter – Fegyő, Tibor – Tüske, Zoltán – Ircing, Pavel 2007. A Morpho-graphemic Approach for the Recognition of Spontaneous Speech in Agglutinative Languages – like Hungarian. *Proc. of Interspeech* 1497–1500.
- Mori, Masahiro 970/2012. The uncanny valley. (*K. F. MacDorman & N. Kageki, Trans.*). *IEEE Robotics & Automation Magazine* 19(2), doi:10.1109/MRA.2012.2192811. 98–100.
- Moulines, Eric – Charpentier, Francis 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communications* 9: 453–467.

- Möbius, Bernd 2000. Corpus-based speech synthesis: methods and challenges. In Sendlmeier, Walter F. – Hess, Wolfgang (eds.): *Speech and Signals – Aspects of Speech Synthesis and Automatic Speech Recognition*. Frankfurt am Main. 79–96.
- Nagy, Péter – Németh, Géza 2016. DNN-Based Duration Modeling for Synthesizing Short Sentences. *Proc. of Speech and Computer (SPECOM 2016)*. Budapest: Springer International Publishing. 254–261.
- Németh Géza – Olasz Gábor 2010. *A magyar beszéd*. 1. Budapest: Akadémiai Kiadó.
- Németh Géza – Olasz Gábor – Fék Márk 2006. Új rendszerű, korpusz alapú gépi szövegfelolvasó fejlesztése és kísérleti eredményei. In *Beszédkutatás* 183–196.
- Olasz Gábor 1989. *Elektronikus beszédelőállítás. A magyar beszéd akusztikája és formánsszintézise*. Budapest: Műszaki Könyvkiadó.
- Olasz, Gábor 1989. MULTIVOX – A flexible text-to-speech system for Hungarian, Finnish, German, Esperanto, Italian and other languages for IBM-PC. *Proc. of Eurospeech '89*. Paris: European Speech Communication Association. 525–529.
- Olasz Gábor 1978. Szintetizált magyar magánhangzók formáns-intenzitás és formáns-sávszélesség értékei. *Magyar Fonetikai Füzetek* 68–77.
- Olasz, Gábor – Németh, Géza 1999. IVR for Banking and Residential Telephone Subscribers Using Stored Messages Combined with a New Number-to-Speech Synthesis Method. In Gardner-Bonneau, D. (ed.): *Human Factors and Voice Interactive Systems*. New York: Kluwer Academic Publishers. 237–256.
- Olasz, Gábor – Németh Géza, – Olasz, Péter – Kiss, Géza – Zainkó, Csaba – Gordos, Géza 2000. Profivox– a Hungarian TTS System for Telecommunications Applications. *International Journal of Speech Technology* 3–4: 201–215.
- Rosenberg, Aaron. G. – Schafer, Ronald. W. – Rabiner, Lawrence. R. 1971. Effects of Smoothing and Quantizing the Parameters of Formant-Coded Voiced Speech. *J. Acoust. Soc. Am.* 50 (6B): 1532–1538.
- Stevens, Kenneth N. – Kasowski, Stanley – Fant, C. Gunnar M. 1953. An electrical analog of the vocal tract. *Journal of the Acoustical Society of America* 25 (4): 734–742. doi:10.1121/1.1907169.
- Tokuda, Keichi – Yoshimura, Takayoshi – Masuko, Takashi – Kobayashi, Takao – Kitamura, Tadashi 2000. Speech parameter generation

- algorithms for HMM-based speech synthesis. *Proc. of ICASSP*. Istanbul, Turkey: IEEE, 2000. 1315–1318.
- Tóth, Bálint Pál – Németh, Géza 2008. Hidden Markov Model Based Speech Synthesis System in Hungarian. *Infocommunications Journal* LXIII (7): 30–34.
- Tóth, Bálint Pál – Németh, Géza 2010. Improvements of Hungarian Hidden Markov Model-based Text-to-Speech Synthesis. *Acta Cybernetica-Szeged* 19 (4): 715–731.
- Van den Oord, Aaron – Kalchbrenner, N. – Kavukcuoglu, K. 2016. Pixel Recurrent Neural Networks. *arXiv preprint arXiv:1601.06759*.
- Van den Oord, Aaron – Dieleman, Sander et al. 2016. A generative model for raw audio. *arXiv preprint arXiv:1609.03499*.
- Waibel, Alex. – Hanazawa, Toshiyuki – Hinton, Geoffrey – Shikano, Kiyohiro – Lang, Kevin J. 1989. Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37 (3): 328–339.
- Werbos, Paul 1974. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis: Harvard University.
- Zainkó Csaba – Tóth Bálint Pál – Németh Géza 2017. Magyar nyelvű WaveNet kísérletek. *Szegedi Egyetem*. Szeged: XIII. Magyar Számítógépes Nyelvészeti Konferencia. 205–216.
- Zainkó, Csaba – Bartalis, Mátyás – Németh, Géza – Olasz, Gábor 2015. A Polyglot Domain Optimised Text-To-Speech System for Railway Station Announcements. *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*. Dresden: International Speech Communication Association. 1236–1240.
- Zen, Heiga – Senior, Andrew – Schuster, Mike 2013. Statistical Parametric Speech Synthesis Using Deep Neural Networks. *IEEE*. New York: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. 7962–7966.
- Zen, Heiga – Tokuda, Keichi – Black, Allan W. 2009. Statistical parametric speech synthesis. *Speech Communication* 51: 1039–1064.

Laki László János

Mesterséges intelligencia a gépi fordításban

1. Bevezetés

A nyelvtechnológia egyik legfontosabb feladata a nyelvi diverzitás okozta akadályok áthidalása, vagyis a számítógépek alkalmassá tétele különböző nyelvek közti fordítások megvalósítására. Az elmúlt évtizedben az információtechnológia robbanásszerű fejlődése lehetővé tette a számítógépes nyelvészet számára, hogy megoldást nyújtson erre a problémára. A feladat mindig is kiemelt fontosságú volt az emberiség számára, ezért pénzt, erőt és energiát nem kímélve igyekeztek megoldást találni a természetes szövegek számítógéppel történő fordítására. Ennek ellenére az elmúlt több mint hatvan év fejlesztései során nem sikerült megvalósítani a tökéletes mesterséges fordítórendszert, viszont a mesterséges intelligencia egyre szélesebb körű alkalmazásával egyre inkább megjelennek az ember számára elfogadható minőségű fordítások. A számítógépes fordítórendszereket a fordítástámogatás területén sikerült több-kevesebb sikerrel alkalmazni.

A folyamatosan fennálló nehézségek ellenére a napjainkban megjelent legújabb számítógépes erőforrásoknak (videokártya/GPU), valamint a neurálhálózat-alapú gépi tanulásnak köszönhetően jelentős változás figyelhető meg a legtöbb tudományág területén. A mesterséges intelligencia viszonylag nagy pontossággal képes felismerni és megkülönböztetni a tárgyakat és a személyeket képekről, valamint a térfigyelő kamerák képeiről. Napjainkban győzte le a mesterséges intelligencia az emberiség legnehezebb játékának tartott GO-ban az aktuális

nagymestert,¹ valamint a sakkban is olyan stratégiákat talál ki, melyeket a sakknagymesterek próbálnak értelmezni és megtanulni tőle.²

Ez a jelenség és tendencia a gépi fordításban is megfigyelhető. A legújabb mélytanulós módszereknek köszönhetően a gépi fordítás eredményessége megduplázódott, és a legfontosabb nyelvpárok között (angol–spanyol, angol–német, angol–francia) sosem látott minőséget produkál. Ez még nem jelenti azt, hogy a humán fordítók elveszítik munkájukat, csupán a jobb minőségű előfordító rendszereknek köszönhetően sokkal eredményesebben képesek dolgozni.

Jelen írásomban bemutatom a gépi fordítás történetét, majd ismertetem napjaink piacvezető technológiájának számító neurális hálózat-alapú gépi fordítórendszer működését. Végül bemutatom ezen rendszerek néhány érdekes alkalmazási területét, úgymint a képhez történő szövegenerálás vagy a hang leiratozása.

2. A gépi fordítás története

A nyelvtechnológia egyik jelentős területe a soknyelvűség támogatása, amire a napjainkban igen hangsúlyos globalizációs törekvések miatt egyre növekvő igény van. Ebben nyújt támogatást a gépi fordítás, melynek módszerei nem csupán a nyelvek közötti transzformáció megvalósításáról szólnak, hanem tartalmazzák a szövegek elő- és utófeldolgozását, valamint a fordítások minőségének előzetes becslését, illetve azok kiértékelését is. Az elmúlt közel hatvan év során számos megközelítés született a természetes nyelvek közötti fordítás megoldására, amelyek közül a legfontosabb állomásokat mutatom be. A gépi fordítás fejlődését nagymértékben befolyásolja a számítógépek teljesítményének növekedése is, hiszen a személyi számítógépek fejlesztése során

¹ Silver–Schrittwieser és mtsai 2017.

² Silver–Hubert és mtsai 2017.

megfigyelhető technológiai (generációs) váltásokat rövid idő eltelével követték a gépi fordítás területén a jelentős módszertani változtatások.

2.1. A gépi fordítás korai története

A gépi fordítás tudománya egyidős az első számítógépek megjelenésével, és mind a mai napig a számítógépes nyelvészet egyik leginkább kutatott területe. Az egyik legelső fordítórendszernek Alan Turing és csapata által létrehozott elektromechanikus rendszer tekinthető, mely segítségével sikerült feltörni a II. világháború során a németek legfejlettebb titkosítási algoritmusát, az úgynevezett Enigma-kódot.

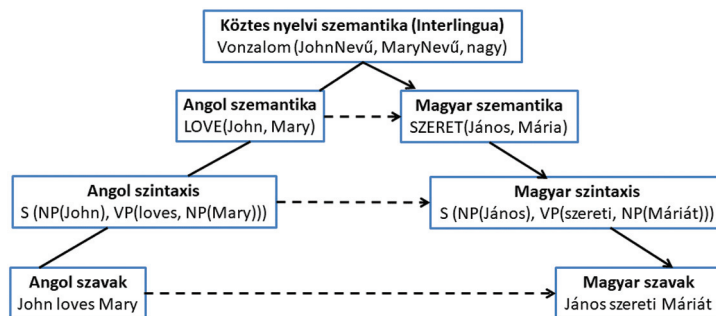
A világháború után a gépi fordítás történetének mozgatórugója a hidegháború volt. Az amerikai kormány jelentős mértékben támogatta a kutatókat, hogy megoldást találjanak az angol és az orosz nyelvek közötti automatikus fordításra. A támogatás eredményeképp 1954-ben a Georgetown Egyetem és az IBM közösen mutatta be az első gépi fordítórendszert. Az elért eredmények alapján a kutatók jelentős része úgy érezte, hogy a gépi fordítás pár éven belül megoldott probléma lesz. Azonban a kezdeti sikerek és a hatalmas anyagi támogatások ellenére sem történt meg a várva várt áttörés. Az amerikai kormány vizsgálóbizottságot hívott össze, melynek feladata a gépi fordítás hatékonyságának kiértékelése volt. Így 1966-ban elkészült az ALPACH-jelentés, melyben nagyon lesújtó képet festettek a számítógép általi fordítás jelenéről és jövőjéről. A jelentés hatására töredékére esett vissza a fordításra szánt állami támogatások összege. A megcsappant támogatások ellenére a fejlesztések nem álltak le, így sorra jöttek létre az első üzleti megoldások; először 1976-ban a Météo és a Systran, majd 1980-ban a Logo, a METAL, valamint a Trados rendszerei.

2.2. A szabályalapú gépi fordítórendszer

A 80-as, 90-es években a számítógépek teljesítménye nagymértékben javult a korábbi évekhez képest. Egyrészt alkalmasak lettek nagyobb méretű szótárak tárolására és használatára, valamint képessé váltak komolyabb mennyiségű szöveges tartalom feldolgozására, másrészt pedig a személyi számítógép megjelenésével a fejlesztőknek lehetőségük nyílt komolyabb programkódok megírására is. Ebben az időben jelentek meg a gépi fordítás támogatására alkalmazott első számítógépes nyelvfeldolgozó rendszerek, melyek integrációjával megalkották az első szabályalapú gépi fordítórendszereket is.

A szabályalapú gépi fordítórendszer (RBMT – Rule-Based Machine Translation) alapötlete, hogy a fordítandó szövegből kinyerhető legtöbb információt használja fel a fordítás során. A legegyszerűbb első implementációk az úgynevezett direkt fordítórendszerek voltak. A módszer lényege, hogy a fordítandó szöveget egy szótár alapján szóról szóra fordítja le, majd a megfelelő sorrendbe rendezi. Ennek előnye, hogy viszonylag könnyen megvalósítható. Hátránya azonban, hogy nem képes komplex nyelvtani szerkezetek kezelésére, illetve ebből eredendően csak gyenge fordításminőséget képes elérni. A későbbi, bonyolultabb rendszerek a fordítandó szövegből elemzés segítségével állítanak elő egy köztes reprezentációt, amit előre definiált átviteli szabályok segítségével alakítanak át egy absztrakt célnyelvi reprezentációra. Végül ebből a reprezentációból generálják a célnyelvi szóalakokat. Ezeket a rendszereket az elemzés és generálás mélysége, valamint az átvitel helye alapján osztályozhatjuk, melyet a Vauquois-háromszög szemléltet (1. ábra). Az ábrán látható, hogy minél mélyebb a nyelvi elemzés mértéke a fordítás során, annál közelebb áll egymáshoz a két nyelv reprezentációja, amelyek közt a transzformációt végre kell hajtani. Egy precízen megírt szabályokkal rendelkező szabályalapú gépi fordítórendszer nagy pontosságú fordítást képes előállítani, de az átviteli szabályok létrehozásához elengedhetetlen a jó minőségű szintaktikai és/vagy szemantikai elemző, ami csak nagyon kevés nyelv esetén áll ren-

delkezésére. Továbbá, mivel ezek a szabályok nyelvspecifikusak, minden nyelvpárra külön-külön kell definiálni ezeket, ami megnehezíti a rendszer kiterjesztését újabb nyelvekre.



1. ábra. Vauquois-háromszög³

2.3. Statisztikai gépi fordítás

Az 1990-es évek elejére a digitálisan elérhető párhuzamosan lefordított szövegek száma ugrásszerűen megnőtt ezzel újabb lehetőségeket nyitva a fordítástámogató rendszerek számára. A szabályalapú rendszerekkel párhuzamosan megjelentek az adatmotivált módszerek. Az első ilyen rendszernek a *példaalapú fordítórendszer* számít, melynek alapötlete, hogy az aktuális fordításhoz felhasználja a már korábban lefordított mondatokat. A rendszer egy előre létrehozott fordítómemóriából kiválasztja a fordítandó mondat részeinek eltárolt fordításait, melyek egyesítésével megkapjuk a kívánt fordítást.⁴ Annak ellenére, hogy a rendszer nem tartalmaz komplex, nyelvspecifikus modulokat, fordítási minősége nem sokkal marad el szabályalapú társaitól. Alapvető hiányossága azonban, hogy a fordítómemóriában tárolt

³ Indurkha–Damerau 2010; Koehn 2010.

⁴ Hutchins 2005.

szegmenspárok elemi egységei (például morfémák, szavak, kifejezések stb.) nincsenek összekötve. Ennek következtében hiába tudja a fordítórendszer, hogy a memóriában tárolt forrásnyelvi szegmens melyik részében különbözik a fordítandó szegmenstől, azt viszont nem tudja meghatározni, hogy a célnyelvi oldalon ez melyik szavakra van hatással.

A *statisztikai gépi fordítórendszer* (SMT – Statistical Machine Translation) a példaalapú fordítórendszer általánosított változatának tekinthető, mivel képes javaslatokat tenni a fordítómemóriában nem szereplő szegmensek fordítására is. A statisztikai gépi fordítás alapötlete, hogy a rendszer párhuzamos kétnyelvű tanítóanyag segítségével felügyelt módon tanulja meg a fordításhoz szükséges modelleket. A párhuzamos kétnyelvű korpusz egy olyan, mondatpárokból álló szöveges adathalmaz, melyben a forrásnyelvi mondatokhoz hozzá van rendelve azok célnyelvi fordítása. Az algoritmus könnyű és gyors implementálhatósága, valamint nyelvfüggetlen alkalmazhatósága nagymértékben hozzájárult ahhoz, hogy a módszer napjainkra a legtöbbet hivatkozott gépi fordító architektúra legyen.

Az SMT alapjait az IBM T. J. Watson Research Center munkatársai fektették le,⁵ akik a fordítás feladatát a beszédtechnológiában használatos Shannon-féle zajoscsatorna-modell⁶ segítségével közelítették meg. A későbbi kutatások eredményeként napjainkra a zajoscsatorna-modell kiegészített változatát alkalmazzák, az úgynevezett log-lineáris modellt.⁷ A következőkben bemutatom ezt a két modellt, valamint ezek kapcsolódását a gépi fordításhoz.

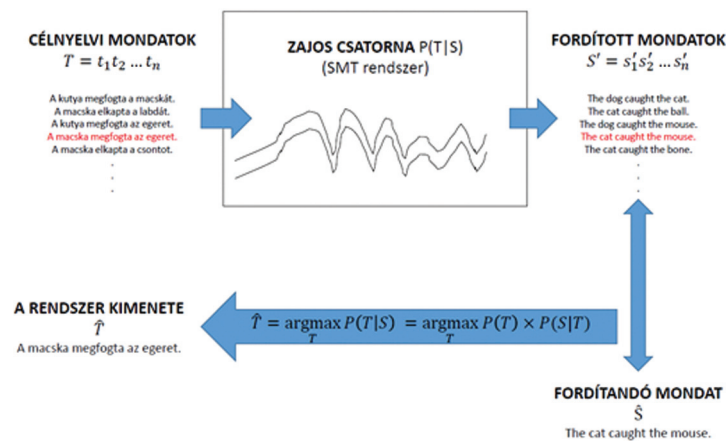
⁵ Brownsé mtsai 1993.

⁶ Shannon 1948a, 1948b.

⁷ Och–Ney 2003, 2004.

2.3.1. Zajoscsatorna-modell

A statisztikai gépi fordítás feladata megfogalmazható a Shannon-féle zajoscsatorna-modell⁸ segítségével, amit a 2. ábra mutat. Az elmélet alapja, hogy a fordítás során az egyetlen biztosan ismert információ a fordítandó szöveg. A fordítás úgy történik, mintha a célnyelvi szövegek halmazát egy zajos csatornán átengedve a csatorna kimenetén összehasonlítanánk a forrásnyelvi szöveggel. Az a célnyelvi mondat lesz a rendszer kimenete, amelyik a csatornán való áthaladás után a legjobban hasonlít a fordítandó mondatra.



2. ábra. Zajoscsatorna-modell

Formálisan az SMT-módszer a fordítás feladatát úgy tekinti, mint a fordítás pontosságát, valamint görbülékenységét reprezentáló modellek kombinációja által elérhető maximális valószínűségi értékhez tartozó szöveg meghatározása. A fordítás fel-

⁸ Shannon 1948a, 1948b.

adata tehát úgy fogalmazható meg, hogy keressük azt a célnyelvi mondatot (\hat{T}), amelyik a célnyelvi mondatok halmazából (T) a legvalószínűbb fordítása a forrásnyelvi mondatnak (S).

$$\hat{T} = \operatorname{argmax}_T P(T|S)$$

A $P(T|S)$ valószínűség azonban közvetlenül nem számolható, viszont önmagában modellezhető részekre bontható. A Bayes-tétel alapján az (1) egyenlet átalakítható a következőképpen:

$$\hat{T} = \operatorname{argmax}_T P(T|S) = \operatorname{argmax}_T \frac{P(S|T)P(T)}{P(S)} = \operatorname{argmax}_T P(S|T)P(T)$$

Mivel T függvényében $P(S)$ konstans, ezért elhagyható. Az így kapott egyenlet két komponensből áll:

- $P(T)$ a nyelvmodell, ami a gördülékenyséért felelős
- $P(T|S)$ a fordítási modell, ami a fordítás pontosságát biztosítja

A két modell kombinációjának maximális értékét a fordítórendszer dekódernek nevezett komponense határozza meg. A (2) egyenlet legfontosabb jellemzője, hogy a fordítás feladatát kiszámítható egységekre bontja. Ráadásul e komponensek becslése egy- és kétnyelvű korpuszok segítségével automatikusan történik.

A zajoscsatorna-modell használata számos megszorítást vezet be, amik korlátozzák a fordítórendszer minőségét. Ilyen megszorítás például, hogy ez a modell egy szónak vagy egy kifejezésnek a fordítását a környező szavaktól függetlennek tekinti, illetve hogy a nyelvmodell csak néhány megelőző szót vesz figyelembe a fordítás során. Sajnos ezek a korlátok jelentősen csökkentik a fordítórendszer minőségét. Emiatt szükségessé vált, hogy a fordításhoz a nyelvmodellen és a fordítási modellen kívül egyéb tudás is felhasználható legyen. Erre ad megoldást a log-lineáris modell, melyet az alábbiakban mutatok be.

2.3.2. *Log-lineáris modell*

A log-lineáris modell a gépi tanulás tudományágának egyik gyakran használt módszere. Lényege, hogy egy feladatot egymástól független jellemzők ($h_i(\chi)$) súlyozott szorzatával (λ_i) írja le. Formálisan a modell a következő:

$$P(x) = \exp \sum_{i=1}^n \lambda_i h_i(x)$$

ahol $h_i(\chi)$ az i . jellemző függvény, míg λ_i a hozzá tartozó súly.⁹ A log-lineáris modell bevezetésével a gépi fordítórendszer tanítása egy újabb lépéssel bővül. Első lépésben be kell tanítani a rendszer modelljeit külön-külön, majd egy úgynevezett paraméter optimalizációs (tuning) lépésben keressük azokat az λ_i értékeket, amivel a fordítórendszer a legjobb minőségű fordítást állítja elő.

Ez a modell a zajoscsatorna-modell kiegészítésének tekinthető, mivel a nyelvmodell és a fordítási modell mellett tetszőleges számú komponenssel bővíthető. Ezek a komponensek lehetővé teszik, hogy a fordítórendszer mélyebb nyelvi (főleg szintaktikai vagy grammatikai) ismeretekkel rendelkezzen, és ezeket felhasználva javítson a fordítások minőségén. A gyakorlatban elterjedt modellek közös jellemzője, hogy mindegyik egyszerű nyelvfüggetlen megszorításokkal igyekszik a fordítási állapotot csökkenteni, melynek köszönhetően sikerült javítani a fordítások minőségét az angol–spanyol, az angol–francia és az angol–német nyelvpárok között. A biztató eredmények ellenére az angol nyelvre kiválóan működő megszorítások a magyar nyelv esetén kártékonynak bizonyultak. A teljesség igénye nélkül nézzünk meg néhány kiegészítő modellt.

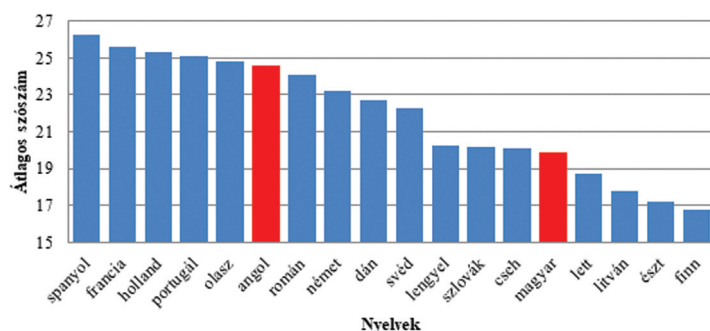
⁹ Och–Ney 2003, 2004.

2.3.2.1. SZÓRENDBELI KÜLÖNBSÉGET BÜNTETŐ MODELL

A szórendbeli különbséget büntető modell (distortion model) feladata a fordítás során a szavak helyes sorrendjének a meghatározása, valamint a keresési gráf méretének a csökkentése. Ez az átrendezési modell azzal a megközelítéssel él, hogy a célnyelvi mondat szavainak sorrendje hasonlít a forrásnyelvi mondat szavainak sorrendjéhez. Ez a feltevés igaz az egymáshoz grammatikailag közel álló nyelvekre, viszont problémát okoz például angol–német nyelvpárok közötti fordítás esetén, ahol a mellékmondat végén szereplő ige fordítását a modell olyan súlyos mértékben bünteti, hogy az SMT-rendszer egyszerűen nem fordítja le azt. Hasonló probléma jelentkezik az angol–magyar vagy a japán–angol fordítás esetében is.

2.3.2.2. A MONDATHOSSZ-HARMONIZÁCIÓS MODELL

A mondathossz-harmonizációs modell azzal a feltételezéssel él, hogy az eredeti mondat és a lefordított mondat szavainak száma hasonló. Ennek megfelelően a modell feladata, hogy kiszűrje a szószámában jelentősen eltérő fordítási javaslatokat. Könnyen belátható, hogy a modell által használt feltételezés túlságosan erős feltétel. A 3. ábra egy példát mutat be, ahol az Európarl



3. ábra. A mondatok átlagos szószáma

korpusz¹⁰ alapján kiszámoltam, hogy átlagosan hány szóból áll egy mondat a korpuszban szereplő nyelvekben. A diagram jól szemlélteti, hogy a különböző típusú nyelvek között jelentős eltérés mutatkozik a mondatok átlagos szószámában, emiatt a mondathossz-harmonizációs modell hangsúlyozott figyelembevétele ronthatja a fordítórendszer minőségét.

2.3.2.3. A LEXIKALIZÁLT ÁTRENDÉZÉSI MODELL

A lexikalizált átrendezési modell (phrase based lexicalized reordering model) lényege, hogy egy kifejezés fordítása után három úton folytatódhat a fordítás folyamata: 1. balról jobbra történő monoton fordítás; 2. a soron következő kifejezés átugrása; 3. a megelőző kifejezés fordítása.

Például az „*I can count the stars visible.*” mondat esetén a *the* szó fordítása után nem a soron következő *stars* szó fordítása következik, hanem az azt következő *visible* szóé, majd ezután visszalép a megelőző szóra. Így kapjuk meg a „*Meg tudom számolni a látható csillagokat.*” fordítást.

Doktori disszertációmban az ezen modellek okozta minőségromlást igyekeztem csökkenteni saját, magyar nyelvre specifikus nyelvfüggetlen modulok létrehozásával. Munkám során sikerült különböző technikákkal (szintaktikai elemzés és generálás, szórendi átrendezések) a magyar és az angol nyelvű mondatokat grammatikailag közelebb hozni egymáshoz, és ezáltal javulást értem el a gépi fordítások minőségében.

2.4. Neurális hálózat-alapú gépi fordítás

Az évtized elejére az SMT-alapú gépi fordítással a kutatóknak nem sikerült jelentősebb minőségjavulást elérni, így indultak el

¹⁰ Koehn 2005.

a kutatások újabb és újabb technikák irányába. Ezekben az években kezdte meg hódító útját a mesterséges neurális hálózatok használata a gépi tanulás területén. A módszer egyáltalán nem tekinthető újnak csupán a számítógépek teljesítményének korlátossága akadályozta ezen technológia használatát. A beszédtechnológiában elért komoly sikereken felbuzdulva a kutatók megvizsgálták a neurális hálózatok alkalmazhatóságát a gépi fordítás feladatának megoldására.

Philipp Koehn a neurális hálózat-alapú gépi fordítás (NMT) fejlődését bemutató könyvének¹¹ bevezetőjében rámutatott arra, hogy ez a technológia már a 80-as, 90-es években is rendelkezésünkre állt, és más tudományterületeken sikerrel alkalmazták. Az 1990-es években a kutatók megpróbálták kiaknázni a neurális hálózatokban rejlő perspektívákat a gépi fordítás területén is. Forcada és Neco¹², valamint Castaño és társai¹³ már 1997-ben létrehoztak olyan neurális hálózat-alapú rendszereket, melyek fordítási minősége megközelítette a létező domináns rendszerek minőségét. Ennek ellenére, mivel az akkoriban rendelkezésre álló erőforrások még nem voltak alkalmasak nagy adathalmazok kezelésére, az egyik ilyen modellt sem tudták megfelelő módon tanítani ahhoz, hogy számottevő eredményeket tudjanak elérni. Ebben az időben ugyanis a rendszerek számítási komplexitása messzemenően meghaladta a korszak számítási erőforrásait, így ezzel a módszerrel lényegében két évtizedig senki sem foglalkozott. Ez alatt a nyugalmi időszak alatt az olyan adatközpontú megoldások, mint például a frázisalapú statisztikai fordítórendszerek, nagy fejlődésen mentek keresztül, és hasznos segítséget nyújtottak számos fordítási probléma megoldásában, illetve növelték a szakfordítók teljesítményét is.

A hagyományos SMT-rendszerek neurális modellekkel történő kiegészítése a neurális módszerek iránti érdeklődés feltáma-

¹¹ Koehn 2017.

¹² Forcada–Neco 1997.

¹³ Castaño és társai 1997.

dását eredményezte. Az áttörést Schwenk¹⁴ hozta, amikor az ily módon felépített hibrid rendszerével jelentős fordításminőség-javulást ért el. Ezek a törekvések azonban a számítási komplexitás miatt csak lassan nyertek teret. Több kutatócsoport is dolgozott a neurálisháló-alapú rendszerek GPU-kon történő tanításának megvalósításán, hiszen a GPU-kban eddig még kiaknázatlan lehetőségek rejlettek.

Az ugrásszerű fejlesztéseknek köszönhetően egy-két év leforgása alatt a gépi fordítás területén is a neurálisháló-alapú kutatások lettek népszerűek a kutatók körében. Összehasonlításképp, az évente megrendezett Conference on Machine Translation (WMT) megmérettetésén 2015-ben még csak egy neurálisháló-alapú fordítórendszert mutattak be, amely versenyképes volt ugyan, de a hagyományos rendszerek még jobbak voltak. 2016-ban már minden nyelvpár-kategóriában a neurálisháló-alapú rendszerek fölényeskedtek, 2017-ben pedig már majdnem az összes benevezett fordítórendszer alapja a neurális hálózat volt. Ez a fejlődés még mindig tart, és az elkövetkezendő években több irányvonal fog megjelenni a rendszerek tökéletesítése céljából.

3. Mesterséges neurálishálózatok bemutatása és alkalmazása

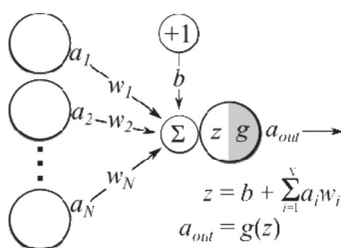
A továbbiakban a neurálishálózat-alapú gépi fordítás módszerével ismerkedhetünk meg.

3.1. Mesterséges neuronok működése

Ahogy azt a 2.3.2. fejezetben említettem, a statisztikai gépi fordítórendszerek a log-lineáris modell segítségével valósítják meg a

¹⁴ Schwenk 2007.

fordítás feladatát. A fordítandó mondatból kinyerhető, valamint a rendszer tanítása során felépített modellekből számolt jellemzők (angolul feature) súlyozott összege segítségével számoljuk ki a rendszer kimenetén megjelenő célnyelvi mondatot. A log-lineáris modellt legegyszerűbben egy hálózattal lehet szemléltetni, ahol a bemeneten a jellemzők, az éleken az azokhoz tartozó súlyok, míg a kimeneten ezek összege szerepel. Ezt a reprezentációt a 4. ábra mutatja. A gyakorlatban ez a neuron bemeneti és kimeneti vektorokból (a_i , a_{out}), a súlyok vektorából (w_i) (és kettő szorzatából tevődik össze).



4. ábra. Mesterséges neuron sematikus rajza¹⁵

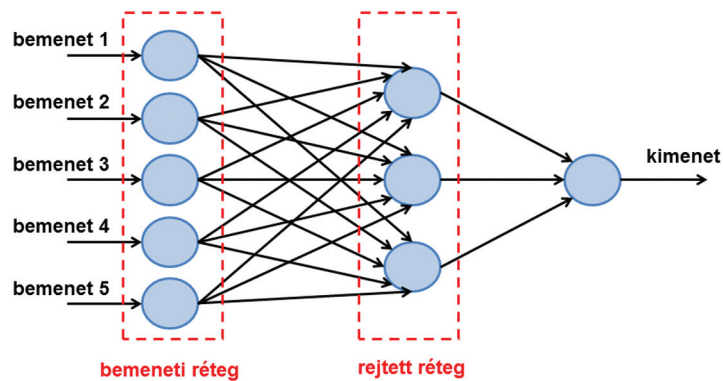
A neurális hálózatok a lineáris modelleket kétféle szemszögből változtatják meg. Az első, hogy a bemeneti értékekből direkt módon számolt kimenetek helyett többrétegű architektúrát alkalmaznak. Ezt a köztes réteget rejtett rétegnek nevezzük, hiszen csak a rendszer bemenetét és kimenetét látjuk, azonban a köztes belső kapcsolatokat nem. A neurális hálózatok másik erőssége a log-lineáris modellel szemben az úgynevezett nem linearitás bevezetése. Ez azt jelenti, hogy a log-lineáris modell egyenletén végrehajtunk egy nem lineáris transzformációt, aminek a következő hatásai vannak a rendszerre: 1. képes egy neuront teljesen kikapcsolni; 2. lehetséges egy neuront félig bekapcsolt állapotba hozni; valamint 3. lehetséges a neuront bekapcsolt állapotban

¹⁵ <https://theclevermachine.files.wordpress.com/2014/09/perceptron2.png>

tartani. Ezeknek a tulajdonságoknak köszönhetően a rejtett réteg képes megtalálni a fontos jellemzőket, és csökkenteni a nem releváns vagy ártó jellemzők hatását.

3.2. Előrecsatolt és rekurrens neurális hálózatok

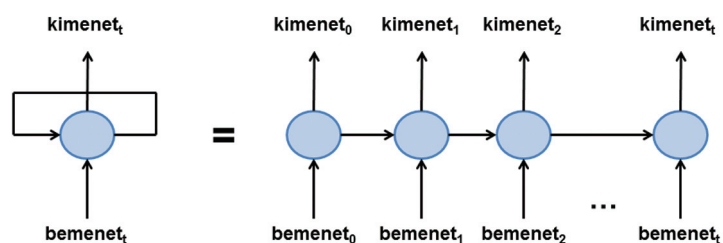
Az 5. ábra egy rejtett réteggel rendelkező neurális hálózatot ábrázol, aminek öt független bemeneti és egy darab kimeneti értéke van. Ezt az architektúrát előrecsatolt neurális hálózatnak nevezük. Az ilyen típusú hálózatok rendkívül jól teljesítenek osztályozási feladatok elvégzésére, mint például a képről tárgy-, arc- vagy személyfelismerés. Mivel az előrecsatolt neurális hálózatok paramétereit és jellemzőit (tehát a különböző rétegek neuronjainak a számát) a tanításuk elején rögzíteni kell, emiatt az ilyen típusú hálózatok nem alkalmasak folytonos hang, illetve szöveg feldolgozására, mivel ezeknél a feladatoknál nem ismerjük előre a bemeneti elemek számát.



5. ábra. Előrecsatolt neurális hálózat egy rejtett réteggel

Erre a feladatra jelent megoldást a rekurrens neurális hálózatok (RNN – Recurrent Neural Network) használata. A rekurrens

hálózat lényege, hogy a mesterséges neuron kimenetét visszacsatoljuk ugyanannak a neuronnak a bemenetére. A szövegfeldolgozás során a vizsgált szöveg szavai egymás után lesznek a neuron bemenetei. A kimenet visszacsatolásának köszönhetően minden egyes szó feldolgozására hatással vannak az őt megelőző szavak is. Az 6. ábra egy önmagába visszacsatolt neuront ábrázol. A rekurrens neurális hálózatokat szokás kifejtett módon is ábrázolni, ahol szemléletesebben látszik a tényleges működése.



6. ábra. Rekurrens neurális hálózat sematikus rajza

A neurálishálózat-alapú gépi fordítás során a fordítandó mondat feldolgozása során ilyen rekurrens neurális hálózatokat használunk. Mielőtt részletesebben leírom a napjaink legjobb minőségű NMT rendszerét az úgynevezett enkóder-dekóder alapú NMT-t, be kell mutatnom a szavak kódolására használatos szóbeágyazás modell nevű technikát is.

3.3. Szóbeágyazás-modell

A nyelvtechnológia egyik fontos célja, hogy a dokumentumokat, a szövegeket, illetve a szavakat olyan reprezentációban tudjuk eltárolni, amely a gép számára könnyen feldolgozható. Fontos, hogy ez a reprezentáció magában foglalja a lehető legtöbb információt, és ezeket a lehető legkisebb helyen tárolja. A kérdés az, hogy melyik reprezentáció képes a szavak felszíni alakja mellett



7. ábra. Két hasonló jelentésű mondat szavainak leképezése a szóbeágyazási modell segítségével

azok morfoszintaktikai, szintaktikai vagy akár szemantikai információit is megőrizni.¹⁶

Napjainkban a hagyományos disztribúciós szemantikai modellek helyett a számítógép számára sokkal kényelmesebb folytonos vektoros reprezentációk alkalmazása terjedt el.¹⁷ A módszer lényege, hogy a rendszerben a lexikai elemek egy valós sokdimenziós vektortér egyes pontjai, melyek úgy helyezkednek el az adott térben, hogy az egymáshoz szemantikailag hasonló szavak egymáshoz közel, míg a különböző szavak egymástól távol kerülnek egymástól. A reprezentáció további előnye, hogy értelmezhető két szóvektor különbsége, ami a szavak jelentésbeli hasonlóságát vizsgálja, valamint a szóvektorok összege is, ami a közös jelentésüket tárolja.

A szóbeágyazás-modell tanítása a statisztikai gépi fordító-rendszerekhez hasonlóan nagyméretű szöveges korpusz segítségével történik oly módon, hogy a tanítóanyag minden egyes szavát, valamint a vizsgált szó környezetében lévő szavakat egy rejtett rétegű előrecsatolt neurális hálózat bemenetére helyezzük. A tanítás végén a szavak vektoros reprezentációja a rejtett réteg vektorállapotaiból olvasható ki. A 7. ábrán két azonos jelentésű mondat, szavainak vektoros ábrázolása olvasható. Látható, hogy

¹⁶ Siklósi–Novák 2016a.

¹⁷ Mikolov–Chen és mtsai 2013; Mikolov–Yih és mtsai 2013.

a hasonló jelentésű valamint a szinonima szavak a vektortérben közel helyezkednek el (például *finom* és az *ízletes*), míg a különböző jelentésűek távolabb helyezkednek el (például *fővárosi* és a *fagylalt*).

Az ebben a fejezetben bemutatott példamondat szavait a Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Kar Nyelvtechnológiai Kutatócsoportja által létrehozott szóbeágyazási-modell¹⁸ segítségével helyeztem el a vektortérben. A rendszernek létezik egy lekérdező felülete, amely segítségével bármelyik tetszőleges szóhoz megtudhatjuk a hozzá legközelebbi szavak listáját. A 8. ábrán az előző ábrán (7. ábrán) szereplő példamondat szavai láthatóak, valamint a vektortérben a hozzájuk legközelebbi elemek szerepelnek. Például az első táblázatból kiderül, hogy a *fagylalt* szóhoz a *jégkrém*, a *fagyi*, valamint a *sütemény* szavak hasonlítanak legjobban, de a *szörp* és a *lekvár* szavak is viszonylag hasonló környezetben szerepelnek, mint a *fagylalt* szó.

4. Neurálishálózat-alapú gépi fordítás működése

Ahogy azt a 2.4. fejezetben megemlítettem, napjainkra a kutatók több különböző felépítésű NMT-implementációt hoztak létre, amelyek kisebb-nagyobb eltéréssel, de nagyjából hasonló minőségben végzik a fordítást a különböző nyelvpárok között. Ezen megvalósítások közül a leggyakrabban alkalmazott technika az úgynevezett enkóder-dekóder architektúra, ami napjaink legjobb minőségű gépi fordító megvalósításának számít. A módszer lényege, hogy a fordítás feladatát két elkülöníthető egységgel valósítja meg, ahol az első rész a fordítandó szöveget alakítja át egy szóbeágyazás-alapú reprezentációba, míg a második rész ebből a köztes állapotból generálja ki a célnyelvi szavak felszíni alakját. A módszert Kalchbrenner és Blunsom¹⁹ alkalmazták először az

¹⁸ Siklósi–Novák 2016b.

¹⁹ Kalchbrenner–Blunsom 2013.

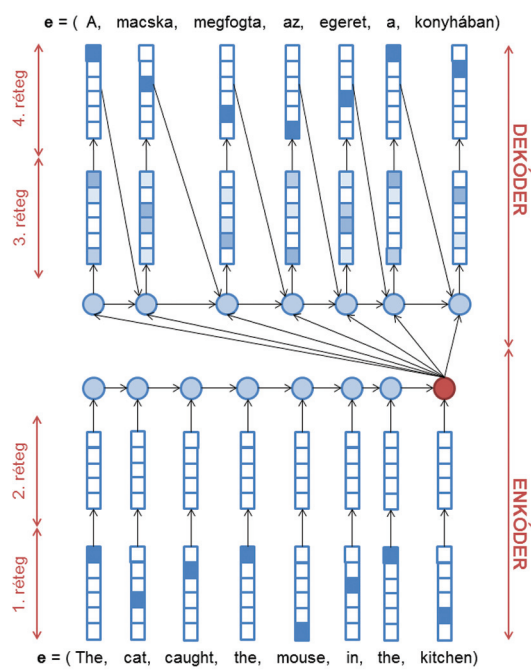
0	fogyólt	1	9133	0	fővárosi	1	129437	0	ízletes	1	18182
1	jéghérm	0.8835	4488	1	kerülőb	0.7771	77525	1	zamatos	0.8263	4732
2	fagyó	0.8783	17395	2	budapesti	0.7759	306934	2	lakatos	0.8262	2812
3	sülemény	0.8597	36852	3	zugói	0.7585	6819	3	fincsiklandó	0.8040	2584
4	desszert	0.8337	18511	4	pécsi	0.7557	128393	4	fincsiklandozó	0.7790	1710
5	csokoládé	0.8295	29077	5	miskolci	0.7473	47096	5	gustuzsos	0.7487	3726
6	édesség	0.8200	38281	6	kaposvári	0.7456	36611	6	ízletes	0.7433	302
7	szendvics	0.8194	23288	7	városi	0.7431	214867	7	omlós	0.7410	3160
8	szörp	0.8109	5054	8	jászévházi	0.7398	5117	8	évszgygyfejlesztő	0.7390	2151
9	likvőr	0.8095	16148	9	szombathelyi	0.7331	44923	9	ehető	0.7362	10577
10	teaszemély	0.8055	619	10	székesfehérvári	0.7221	23085	10	lelős	0.7311	2105

8. ábra. A fagyólat, fővárosi és ízletes szavakhoz legjobban hasonlító szavak listája

Oxfordi Egyetemen, majd idővel számos kutatócsoport²⁰ folytatta a rendszer tökéletesítését. A 9. ábra mutatja be az enkóder-dekóder architektúra sematikus rajzát.

4.1. Az enkóder működése és felépítése

Az enkóder feladata a forrásmondat elhelyezése a megfelelő vektoros reprezentációba, amit úgy valósít meg, hogy a mondat szavait egymás után megmutatja az enkóder belsejében található rekurrens neurális hálónak. A folyamatot a 9. ábra mutatja be,

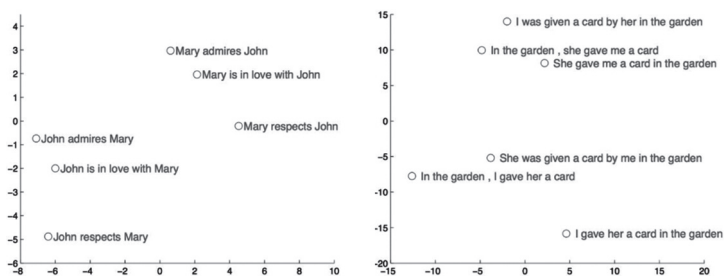


9. ábra. Enkóder-dekóder architektúrájú gépi fordítórendszer sematikus rajza

²⁰ Bahdanau és mtsai 2014; Cho és mtsai 2014.

ahol megfigyelhető a forrásoldali szavak vektoros ábrázolása. Az 1. rétegben a szavakat az úgynevezett „one-hot” reprezentációban tároljuk, ami azt jelenti, hogy a bemeneti szóképlet méretű vektorban egy érték kivételével mindegyik 0 értékű. Ennek a reprezentációnak az előnye, hogy könnyen implementálható, viszont a szavak felszíni alakján kívül semmilyen többletinformációt nem tartalmaznak. Emiatt van szükség a második rétegre, ami a 3.3. fejezetben bemutatott szóbeágyazási modellt valósítja meg.

Miután a mondat összes szava feldolgozásra került, az RNN-neuronban lényegében a forrásmondat vektoros reprezentációja szerepel (a 9. ábrán a piros karikával rajzolt *node* jelöli). Ezzel a módszerrel a szöveget leképeztük egy 300-500 dimenziós vektorba, amik a szóbeágyazási modellhez hasonló tulajdonságokkal rendelkeznek. Erre láthatunk példát a 9. ábrán, ahol a hasonló jelentésű angol és magyar mondatok egymáshoz közel helyezkednek el a térben, míg a különbözőek távolabbi csoportokat alkotnak. Ennek köszönhetően képes ez az NMT-architektúra fordítást végezni, mivel a folytonos vektorreprezentáció hatására a különböző nyelvű mondatok azonos vektortérbe kerülnek.



10. ábra. Példa az NMT-rendszer belsejében szereplő vektoros mondat ábrázolásra²¹

²¹ <https://devblogs.nvidia.com/parallelforall/introduction-neural-machine-translation-gpus-part-2/>

4.2. A dekóder működése és felépítése

A fordítórendszer második fő komponense a dekóder, ami a forrásnyelvi mondat vektoros reprezentációja alapján egy RNN-hálózat segítségével állítja elő a célnyelvi mondatot. Az RNN-hálózat az előző szó felszíni alakja, a fordítandó mondat vektoros reprezentációja, valamint az eddigi fordítások önmagába visszacsatolt állapota segítségével rangsorolja a célnyelvi szótár szavait, hogy melyik szó következik legnagyobb valószínűséggel a fordítás során.

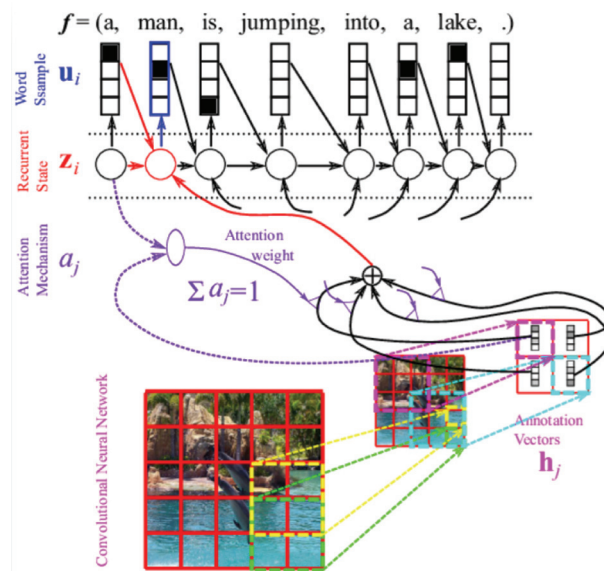
Ezen a ponton merül fel az a kérdés, hogy egy darab 300-500 dimenziós vektor valóban alkalmas arra, hogy tetszőleges hosszúságú mondatnak képes legyen eltárolni oly módon a jelentését, hogy abból egyértelműen előállítható legyen a célnyelvi mondat fordítása? A kérdéssel foglalkozó kutatások²² erre azt a választ hozták, hogy ebben a formájában ez az architektúra nem képes erre. A megoldást Bahdanau és társai²³ mutatták be 2014-ben. Munkájuk során az enkóder-dekóder architektúrát kiegészítették egy figyelmi modellnek nevezett neurális hálózattal, amely a dekódolást végző RNN-neuron bemenetére az aktuális szó fordításához szükséges forrásnyelvi szavakat is figyelembe veszik.

5. Érdekességek a technológia további alkalmazhatóságára

A neurálishálózat-alapú rendszerek óriási előnye, hogy nemcsak a nyelvek szintjére korlátozódnak. Ez azt jelenti, hogy bármilyen adattípust képesek kezelni bemenetként, amennyiben létezik annak megfelelő folytonos vektoriális reprezentációja. Ilyen lehet akár a hang, a videó, a kép stb.

²² Bahdanau és mtsai 2014; Cho és mtsai 2014; Sutskever és mtsai 2014.

²³ Bahdanau és mtsai 2014.

11. ábra. Példa a képleírás generátor működésére²⁴

Nemrégiben a Montreali Egyetemen és a Torontói Egyetemen végzett kutatások bebizonyították, hogy lehetséges olyan figyelemalapú enkóder-dekóder modell kialakítása, amely egy kép alapján szöveges formátumban leírást készít annak tartalmáról. A képet az enkóder konvolúciós neurális hálózattal történő helyettesítésével írja le, majd a figyelmi modell segítségével állítják elő a leírásnak számító mondatot. Ezt a folyamatot mutatja be a 10. ábrán. Hasonló munkákat²⁵ is bemutatott.

Tovább feszegették a határokat Li és munkatársai, amikor 2015-ben hasonló figyelemalapú rendszerekkel kísérleteztek videókra. A video leírásához a dekóder a videóból kinyert ideiglenes struktúrákat használta fel. A közelmúltban több kutatás

²⁴ <https://devblogs.nvidia.com/paralleforall/introduction-neural-machine-translation-gpus-part-3/>

²⁵ Donahue és mtsai 2014; Fang és mtsai 2014; Karpathy–Li 2014.

foglalkozott a videók enkóder-dekóder architektúra segítségével készített leírásával.²⁶ A 12. ábra egy példát mutat, ahol egy főzéssel kapcsolatos videofelvétel képkockái alapján készül el a következő leírás angolul: „Valaki olajban halat süt.” Ezt a négy képkockából kinyert információk alapján állította össze.



12. ábra. Példa filmfelirat generálására²⁷

Végül Chan és munkatársai²⁸ kutatásuk során létrehoztak egy hangból leíratot előállító rendszert. A beszédtechnológiából ismert eljárások segítségével állították elő a hangfelvétel vektoros reprezentációját, amiből karakteralapú enkóder-dekóder hálózat segítségével generálták ki a szöveges kimeneti formátumot.

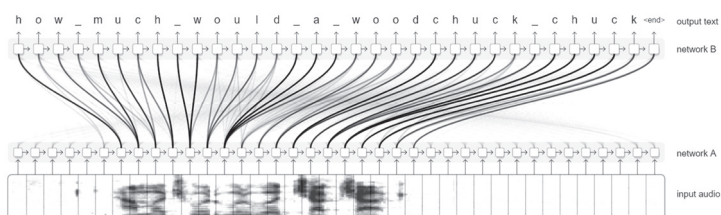
²⁶ Li és mtsai 2015; Venugopalan és mtsai 2015.

²⁷ <https://devblogs.nvidia.com/paralleforall/introduction-neural-machine-translation-gpus-part-3/>

²⁸ Chan és mtsai 2015.

6. Összefoglaló

A Magyar Tudomány Ünnepe keretében elmondott előadásomban bemutattam a mesterséges intelligencia alkalmazását a gépi fordítás területén. Napjainkban a mesterséges intelligencia térnyerésével jelentős változások figyelhetők meg a legtöbb tudományág területén, és ez nincs másként a nyelvtudományokkal sem. A bemutatott új technológiák és architektúrák alkalmazásával sikerült áttörést elérni az eddigi piacvezető kifejezésalapú gépi fordítási módszerekkel szemben és nagymértékű minőségjavulást elérni, még az olyan nehéz nyelvpárok esetén is, mint az angol és a magyar. Azt most nehéz megmondani, hogy mire lesznek képesek a számítógépek a jövőben, és hogy bármikor is sikerül-e létrehozni a tökéletes fordítórendszert, de a tudomány számára nincs lehetetlen.



13. ábra. Példa a hangból szöveg generálásra

Irodalom

- Bahdanau, Dzmitry – Cho, Kyunghyun – Bengio, Yoshua 2014. Neural Machine Translation by Jointly Learning to Align and Translate. *Computing Research Repository (CoRR)*, abs/1409.0473.
- Brown, Peter F. – Pietra, Vincent J. Della – Pietra, Stephen A. Della – Mercer, Robert L. 1993. The Mathematics of Statistical Machine Translation: Parameter Estimation, *Computational Linguistics*, 19 (2): 263–311.

- Castaño, M. – Asunción, Casacuberta, Francisco – Vidal, Enrique 1997. *Machine Translation using Neural Networks and Finite-State Models* 160–167.
- Chan, William – Jaitly, Navdeep, Le – Quoc V. – Vinyals, Oriol 2016. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition, in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 4960–4964.
- Cho, KyungHyun – Merriënboer, Bart van – Bahdanau, Dzmitry – Bengio, Yoshua 2014. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches, *CoRR*, abs/1409.1259.
- Donahue, Jeff – Hendricks, Lisa – Anne, Guadarrama, Sergio – Rohrbach, Marcus – Venugopalan, Subhashini – Saenko, Kate – Darrell, Trevor 2014. Long-term Recurrent Convolutional Networks for Visual Recognition and Description, *CoRR*, abs/1411.4389.
- Fang, Jing – Nevin, Philip – Kairys, Visvaldas – Venclovas, Česlovas – Engen, John R. – Beuning, Penny J. 2014. Conformational Analysis of Processivity Clamps in Solution Demonstrates that Tertiary Structure Does Not Correlate with Protein Dynamics. *Structure* 22 (4): 572–581.
- Forcada, Mikel L. – Neco, Ramón P. 1997. Recursive hetero-associative memories for translation, in Mira, José, Moreno-Díaz, Roberto, and Cabestany, Joan Eds., *Biological and Artificial Computation: From Neuroscience to Technology: International Work-Conference on Artificial and Natural Neural Networks, IWANN'97 Lanzarote, Canary Islands, Spain, June 4–6, 1997 Proceedings*, Berlin, Heidelberg: Springer. 453–462.
- Hutchins, John 2005. Towards a Definition of Example-Based Machine Translation. In *Proceedings of Workshop on Example-Based Machine Translation, MT Summit X*, Phuket, Thailand. 63–70.
- Indurkha, Nitin – Damerau, Fred J. 2010. *Handbook of Natural Language Processing*, 2nd ed. Boca Raton, FL, Chapman & Hall/CRC.
- Kalchbrenner, Nal – Blunsom, Phil 2013. Recurrent Continuous Translation Models, in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Seattle: Association for Computational Linguistics.
- Karpathy, Andrej – Fei-Fei, Li 2015. Deep Visual-Semantic Alignments for Generating Image Descriptions, in *The IEEE Confe-*

- rence on Computer Vision and Pattern Recognition (CVPR) Boston, USA.
- Koehn, Philipp 2005. Europarl: A Parallel Corpus for Statistical Machine Translation, in *Conference Proceedings: the tenth Machine Translation Summit*. Phuket, Thailand: AAMT. 79–86.
- Koehn, Philipp 2010. *Statistical Machine Translation*. New York: Cambridge University Press.
- Koehn, Philipp 2017. Neural Machine Translation, *CoRR*, abs/1709.07809.
- Li, Yao – Torabi, Atousa – Cho, Kyunghyun – Ballas, Nicolas – Pal, Christopher – Larochelle, Hugo – Courville, Aaron 2015. Describing videos by exploiting temporal structure. *Proceedings of the IEEE international conference on computer vision*. 4507–4515.
- Mikolov, Tomas – Chen, Kai – Corrado, Greg – Dean, Jeffrey 2013. Efficient Estimation of Word Representations in Vector Space, in *Proceedings of Workshop at ICLR Scottsdale, AZ*.
- Mikolov, Tomas – Yih, Wen-tau – Zweig, Geoffrey 2013. Linguistic Regularities in Continuous Space Word Representations. *HLT-NAACL*. 746–751.
- Och, Franz Josef – Ney, Hermann 2003. A Systematic Comparison of Various Statistical Alignment Models. *Computational Linguistics* 29 (1): 19–51.
- Och, Franz Josef – Ney, Hermann 2004. The Alignment Template Approach to Statistical Machine Translation. *Computational Linguistics* 30 (4): 417–449.
- Schwenk, Holger (2007) Continuous Space Language Models. *Comput. Speech Lang* 21 (3): 492–518.
- Shannon, Claude Elwood 1948a. A mathematical theory of communication. *Bell System Technical Journal* 27 (July): 379–423.
- Shannon, Claude Elwood 1948b. A mathematical theory of communication. *Bell System Technical Journal* 27 (October): 623–656.
- Siklósi Borbála – Novák Attila 2016a. Beágyázási modellek alkalmazása lexikai kategorizációs feladatokra. In Tanács Attila – Varga Viktor – Vincze Veronika (eds.): *XII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2016)*. Szeged: Szegedi Tudományegyetem. 3–14.
- Siklósi Borbála – Novák Attila 2016. Közeli rokonunk, az autó. In Tanács Attila – Varga Viktor – Vincze Veronika (eds.): *XII. Magyar*

Számítógépes Nyelvészeti Konferencia (MSZNY 2016). Szeged: Szegedi Tudományegyetem. 27–36.

- Silver, David – Hubert, Thomas – Schrittwieser, J. – Antonoglou, I. – Lai, M. – Guez, A. – Lanctot, M. – Sifre, L. – Kumaran, D. – Graepel, T. – Lillicrap, T. – Simonyan, K. – Hassabis, D. 2017. Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. *ArXiv e-prints*.
- Silver, David – Schrittwieser, Julian – Simonyan, Karen – Antonoglou, Ioannis – Huang, Aja – Guez, Arthur – Hubert, Thomas – Baker, Lucas – Lai, Matthew – Bolton, Adrian – Chen, Yutian – Lillicrap, Timothy – Hui, Fan – Sifre, Laurent – van den Driessche, George – Graepel, Thore – Hassabis, Demis 2017. Mastering the game of Go without human knowledge. *Nature* 550: 354–359.
- Sutskever, Ilya – Vinyals, Oriol – Le, Quoc V. 2014. Sequence to Sequence Learning with Neural Networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal: MIT Press. 3104–3112.
- Venugopalan, Subhashini – Rohrbach, Marcus – Donahue, Jeff – Mooney, Raymond J. – Darrell, Trevor – Saenko, Kate 2015. Sequence to Sequence – Video to Text. *CoRR*, abs/1505.00487.

Szűts Zoltán

Irodalmi Nobel-díjas mesterséges intelligencia?

A gépversektől a Big Data rövid történetekig

1. Bevezetés

Miközben a bölcsészettudomány (és irodalomtudomány) figyelmé a számítógépek segítségével létrehozott alapvetően multimedialis, hipertextuális, a jelenben már online vagy éppen többszerzős szépirodalmi szövegek felé fordul, a számítógépek által generált irodalmi szöveg olyan specifikus jelenség, mely ritkán szerepel a diskurzusban. Ilyenkor az alkotási folyamat során egy számítógépes program a felhasználó utasítására irodalmi alkotásokat hoz létre. Az így generált szöveg erősen alá vannak rendelve az őt létrehozó szoftver szabályainak, poétikáját, esztétikáját, stílusát, és gyakran tartalmát is alapvetően a rendszer programozója, és nem kezelője – látszólagos alkotója – határozza meg. Az alkotó – szerző – szerepe körül kialakult konfúzió érthető. Ahhoz azonban, hogy némileg tisztázzuk a szerzőség kapcsán kialakult helyzetet, először történeti dimenzióba kell helyoznünk a számítógépek által írt alkotásokat.

Némi feloldást igényel a tanulmányunk címe is. Munkánkban ugyanis arra keressük a választ, hogy a mind fejlettebb mesterséges intelligencia, és a hozzá szorosan kapcsolódó Big Data módszere segítségével írott irodalmi szövegek milyen esztétikai értékeket, jelentéseket hordoznak, másrészt pedig befogadói stratégiákat kívánnak. Tanulmányunk a jelenség kontextusba helyezésével kezdődik, így helyet kap benne a gépversek témája is.

2. A számítógépek által írt szövegek taxonómiája

Az első, számítógéppel tömegesen produkált irodalmi alkotások, az 1960-as években népszerűvé vált gépversek nem tartoznak azonos kategóriába azon művekkel, melyek létrehozását a 2000-es évek közepétől a webkettes környezetben a szerzők indukálják számítógép segítségével, és még inkább eltérnek a Big Data kontextusában a jelenben született alkotásoktól.

Ha a felsorolást a számítógéppel generált költészettel kezdjük, akkor elmondhatjuk, hogy ezen szövegek már nem analóg formában, hanem digitálisan rögzítettek, valamint létrejöttükben szerepet játszik a számítógép mint eszköz és a szoftver mint kontextus. Nagyon fontos azonban, hogy ebben az esetben a számítógép nem valódi szerző.

A számítógépek által írt szövegek elméleti kérdéseit először J. C. R. Licklider vetette fel, akinek célja volt, hogy megtalálja a kapcsolódási pontot az emberek és gépek között (ő ezt akkor és későbbi munkáiban is szimbiózisnak nevezte), hiszen a gépek partnerré váltak a probléma megoldásban – témánk esetében az alkotásban is. Ez a szimbiózis már egy ember–gép alkotta egységet feltételezett, melyet az említett *Man-Computer Symbiosis* című művében¹ írt körül számos ötlet formájában, melyek végül egyetlen következtetésre vezettek: az emberek és a társulás „elektronikus tagjai” között végül kooperatív döntéshozási rendszer jön majd létre.

Az első, számítógép által írt szöveg 1959-ben jelent meg, amikor Théo Lutz első számítógéppel készült verseit publikálta,² majd őt követték az 1960-as évek elején Jules, Nake és Noll komputergrafikai kísérletei.³

Fontos rámutatnunk, hogy a számítógép környezetében a szövegek mindig is diszkriminációt szenvedtek, szemben például a vizualitással. Bár a programozás során a szoftverek szövegesen

¹ Licklider 1960.

² Lutz 1959.

³ Szombathy 1977.

vannak rögzítve, a számítógépek által generált produktumok nagy része már a képekre épít. Az új technológiák is ezt az irányt követik, elég csak a virtuális vagy augmentált valóságra gondolni, ahol a tárgyi világra alapvetően interaktív képek és nem szövegek rétegződnek. Szombathy Bálint is felhívja a jelenségre a figyelmet, szerinte a gépileg előállított zeneművek és grafikai alkotásokkal szemben a komputerköltészeti törekvések ennél szerényebb keretek között zajlottak. Szombathy ezt részben azzal magyarázza, hogy a gépileg létrehozott költemények kevés teret kaptak a hagyományos médiumokban, másrészt egyik művészeti ágban sem annyira szilárdak az előítéletek, konzervatívak a nézetek, mint a szépirodalomban. Ez annyiban is igaz, hogy az ilyen típusú költészet kevés fejlődési és önkifejezési potenciállal bír, ugyanis nem a szerző poétikai meggyőződésének, hanem a szoftveres (korlátozott) környezetnek volt alárendelve.

Jelentősen gátolta a gépi líra terjedését és fejlődését – az olvasókhöz való eljutást – annak tudata, hogy a költészet eredendően szemantikai jellegű, tehát a jelentés, az érzelmi töltés, a kinyilatkoztatás dimenziójában mozog, ami a hagyományos poétikai álláspont szerint ellentétben áll az érzéketlenséget, lélektelenséget megtestesítő gépek alaptermészetével.⁴ A gépversek ideáját leginkább azok vetették el, akik a nyelvi alkotómunkát kiegyenlítették a lángész csalhatatlanságának, az isteni küldetésnek a fogalmával. Ellenkezésük vetületében a komputerköltészet egyenlő volt a szentségtöréssel és az istenkáromlással.⁵ Szombathy ugyancsak kiemeli, hogy a 20. század tudósai azt hangsúlyozták, hogy a matematika és a kibernetika nem törekszik a verbális nyelv saját nyelveivel való felcserélésére, hanem arra, hogy a verbális nyelv keretein belül teremtsen meg tudományos elveinek és módszertanának az alkalmazását.

⁴ Az 1960-as években megjelent művek kezdeti recepciójára jellemző a megrökönyödés és az ellenállás, talán leginkább a költők részéről, akik az új, bármely számítógépes felhasználó előtt nyitott kifejezési formában Szombathy szerint „isten küldetésük” megkérdőjelezésének veszélyét látták.

⁵ Szombathy 1977.

Vincent van Mechelen a számítógéppel generált költészetet két csoportra osztotta.⁶ Az első csoport a deriváló számítógéppel generált költészet. Ebben az esetben a számítógép már meglévő szövegeket változtat meg adott utasításkészlet alapján. Ezt az eljárást remixként is értelmezhetjük, mely többek között a posztmodern (irodalom) egyik instrumentuma is. Visszatérve az utasításkészlet kérdésére, az tartalmazhat egy képletet, melynek következtében a kapott szövegben bizonyos minták lesznek fellelhetők, vagy egy véletlengenerátortól kapja az utasításait, ebben az esetben pedig mintáról nem beszélhetünk. A második csoportba ezen felosztás szerint az úgynevezett eredeti számítógéppel generált költészet tartozik. Ennél a típusnál a programozó – felhasználó – szerző egy szintaktikai struktúrát definiál, és megad egy halmazt, melyből a számítógép behelyettesíti a szavakat. Ez lehet egy rövid lista, vagy akár egy lexikon teljes szókészlete is. Ez a fajta költészet tehát nem utánoz már meglévő nyomtatott és betáplált szövegeket, és nem azok lexikai egységeivel dolgozik. Az utóbbi típus továbbfejlesztett változata lesz a Big Data környezetében a mesterséges intelligencia közreműködésével létrejött irodalmi szöveg is.

Míg a deriváló számítógéppel generált költészet esetében a származékszöveg megtartja az eredeti grammatikai struktúráját és formáját, addig ez a második csoport esetében már nem igaz, ami jelentős nehézségek elé állítja a befogadót. Az első esetben tehát a kapott szövegek közelebb állnak a klasszikus, már létező szövegekhez, míg a második esetben szabad versekről beszélhetünk. Mechelen szerint az első esetben a költészet imitatív, még a másodikban innovatív. A jelenben pedig alapvetően az utóbbi dominál.⁷

⁶ Mechelen 1992.

⁷ Szűts 2013.

3. A számítógép kilép a magányból, és elkezd szöveget írni

Egy fontos szemléletváltásra szeretnénk felhívni a figyelmet, a *computer* szó 1945-ig eredetileg ugyanis azt az embert jelölte, aki egyenleteket oldott meg, csak ezután terjedt el a ma ismert használata. A II. világháború idején a számítógépeket (computereket) azért találták fel, hogy bonyolult matematikai problémákat oldjanak meg (compute). Bizonyos mértékben a jelenben is még ugyanezt teszik, azonban a számítógépek ma már mélyen gyökeret vertek a társadalmunkban is, szinte nincs olyan aspektusa a mindennapi életben – a vásárlástól a tanuláshoz, az ismeretszerzéstől a bankolásig, a filmnézéstől a szövegírásig – mely ne támaszkodna informatikai eszközökre. A számítógépek adatokat tárolnak, dolgoznak fel és továbbítanak, kommunikációs hálózatokat működtetnek, és médiatartalmat állítanak elő, játszanak le és osztanak meg. A jelenből szemlélve furcsán hathat, hogy a számítógép ideája hosszú ideig nem kapcsolódott össze a művészetével. Magányos, a művésztől távol maradó gépek sora született, mígnem az 1960-as években felmerült az alkotásra való használatuk. Ez a lehetőség még így is sokáig csupán idea maradt. Theo Lutz szerint annak ellenére, hogy egy program vezérli, a számítógépes processzort eredetileg arra tervezték, hogy az alkalmazott matematika igényeit szolgálja ki, ám hamarosan kiderült, hogy a rendszert sokkal szélesebb körben lehet alkalmazni. A szerző már 1959-ben rámutatott, hogy az eszköznek végtelen felhasználása lehetséges, ennek ellenére sok tudós még azon hamis elképzelés foglya, miszerint a processzorok kizárólag a számokkal kapcsolatos műveletek végrehajtására alkalmasak. Számtalan olyan program létezik már azonban, melyek bebizonyították ezen felfogás túlhaladottságát.⁸

Már az 1950-es évek végén bebizonyosodott, hogy a számítógép a kizárólagosan számolási problémák mellett más tудо-

⁸ Lutz 1959.

mányterületekben és a művészetekben is felhasználásra talált. Ezt azért kell hangsúlyozni, mert míg számunkra a jelenben a számítógép és az irodalom közötti összefonódás már elválaszthatatlannak tűnik (ennek több fokozata van, a szövegszerkesztéstől a hipertextig, esetünkben pedig a gépek által írt szövegekig), addig az első személyi és házi számítógépek először valóban csak számítási problémákat voltak hivatva megoldani. A szöveg tehát a hálózat nélküli gép környezetében is megtalálta a kifejezési formáját, azonban a szerző által feladni kívánt szubjektív üzeneteket nem volt képes közvetíteni.

A számítógéppel generált költészet fejlődésének fontos fejezete Alberta Turner munkájához kötődik. Turner 1965-ben megjelentette a *Returner Re-turned* című, naptárhoz kapcsolódó kötetét. Ebben a szerző 13 2-3 soros versben leírta az év valamennyi hónapját, a januárt kétszer. Az első két hónaphoz angolul a következő szöveg kapcsolódik:

Később 1971-ben Louis Milic betáplálta a vers szavait egy általa létrehozott programba, és a *Returner* címet adta az így kapott költeménynek. Milic nemcsak a szavakat másolta be a generátorba, hanem figyelembe vette a két-három soros struktúráját is. Milic számára kiderült, hogy a szövegnek nemcsak az adott nyelv grammatikai szabályainak kell megfelelnie, hanem szűkíteni kell egy adott téma esetén felhasználható lexémák körét is. Ezt a következőképpen határozza meg: a lexikális kohézió nem feltétlenül eredményez szövegbéli vagy narratív kohéziót is.

4. A véletlen eredménye?

A gépversek megszületése a hálózatok és online média tulajdonságaiból fakadóan bizonyos nézőpont szerint a véletlen eredménye, ahogy a mai ismert világháló is a CERN-ben létrehozott elektronikus faliújság továbbfejlesztett változata, és így véletlen következménye. A jelenség kiváltója, hogy az 1960-as években a kibernetika a beszédet és a nyelvet új, mérhető és szimulálható oldaláról kezdte el vizsgálni. Ennek elméleti alapja, hogy a kuta-

tók úgy látták, a matematikai információelmélet számítási módszerei alkalmazhatatlanok a beszéd területén.

5. Ki a szerző?

Visszatérve a tanulmány elején említett problémára: a számítógéppel generált irodalom elemzési kísérlete során felmerülhet egy új szempont is. Amennyiben a szerző egy programot ír, melynek törvényei szerint a gép létrehozza a verset, akkor a programot író vagy a gép a szöveg szerzője? Egy, a világhálón talált irodalmi szöveggenerátorral készült versnek ki a szerzője? Nevezhető-e a számítógép szerzőnek? Készül-e a számítógép közölni valamit?

Hasonló generátorral kísérletezett Raymond Queneau az 1961-ben Párizsban publikált *Cent mille milliards de poèmes* (Százazermilliárd költemény) című versében, amikor egymás mellé helyezett 10 olyan szonettet, melynek valamennyi első sora kombinálható valamennyi második sossal, egészen az utolsó sorig. Ez a szöveg még nyomtatásban jelent meg, ezt követte a digitális kiadás. Így összesen $10^{140} = 100\,000\,000\,000\,000$, azaz százazermilliárd vers jött létre. Figyelembe véve az irodalomtudományi recepció egyik közismert megállapítását, miszerint a mű az olvasó révén létezik, joggal merül fel a kérdés, hogy létezik-e tehát olyan olvasó, aki képes befogadni ezt a szöveget. Ennek az opusnak az elolvasása x millió évet venne igénybe, ami egy olyan hivatást feltételez, mint a *Cent mille milliards de poèmes* olvasó, egy mesterséget, mely ezer generációkon keresztül, apáról fiúra hagyományozódik, hogy életben tartsa a szöveget. Horváth Iván *Magyarok Babelben* című könyvében megjegyzi, hogy ezek szerint Queneau írta a világirodalom majdnem valamennyi szonettjét.⁹

A legismertebb magyar nyelvű számítógéppel generált költeményt, a *Disztichon Alfát*, 1994-ben, a világháló és hipertext

⁹ Horváth 2000.

korában publikálta Papp Tibor, a szöveg azonban még „lemezbe zárt”. Ez egyrészt arra enged következtetni, hogy a neoavantgárd hagyomány a digitális korban a magányos gépen és gépversben folytatódik, míg a posztmodern törekvések a világhálón, és később a webkettőn, a közösségi média által létrehozott tartalom kontextusában találtak kontinuitásra. Mint már olvashattuk, a *Disztichon Alfa* hordozója egy lemez volt, mely akkor még csak Macintosh gépeken volt használható, a program futtatása után disztichonokat ír a képernyőre. A lemezhez tartozott egy 80 oldalas füzet is, mely magyarázatokat és néhány generált verset is tartalmaz.¹⁰

6. Big Data

A jelenben a Big Data-rendszerek jelentik a legígéretesebb szöveggeneráló rendszereket. A Big Data egyszerre jelent szemantikai, analitikai, adattárolási és hozzáférési kihívást, hiszen napjainkban eddig nem látott nagyságrendű adatok online történő tárolását, számítógépes feldolgozását, a rejtett és váratlan összefüggések megtalálását feltételezi. Ahhoz azonban, hogy mindezen adatmennyiség sikeres feldolgozására és mintázatok felismerésére vállalkozhasson bárki is, szüksége van olyan speciális szoftverekre, melyek a fejlett gépi tanulás rendszerére támaszkodnak. A gépi tanulás (machine learning – M. L.) egyik fontos feltétele a Big Datának, hiszen lehetővé teszi, hogy a számítógépek tanuljanak – mintákat vegyenek észre – anélkül, hogy konkrétan erre programozták volna őket. Ezen esetekben a számítógépeket embereknek kell tanítaniuk, az adatokat kezdetben nekik kell megcímkézniük és osztályozniuk, hogy később, e minta alapján, a gépek önállóan is képesek legyenek tanulni és elemezni az információkat. Abban az esetben például, amikor a gépek a közösségi médiában éppen a legvitatottabb

¹⁰ Szűts 2013.

témákat ismerik fel, ilyen emberi felügyeletet igénylő tanulásra már nincs szükség.¹¹

A Big Data mint tartalom-létrehozó eljárás a jövő egyik lehetséges útja. A Narrative Science's algoritmus például már a strukturált adatok (alapvetően hírek) segítségével képes emberi beavatkozás nélküli narratív történeteket generálni. Így például, a sport területét kiemelve, egy médiaügynökség valamennyi statisztikát birtokolva pillanatokkal a meccs vége után már hosszú elemzést publikálhat online, korábban erre órákat kellett várni, sőt az online média megjelenése után is ez a típusú újságírás a másnap megjelenő nyomtatott sajtót jellemezte¹².

7. A chatbotok mint szövegenerálók

A chatbotok alapvetően olyan számítógépes programok, melyek képesek az emberekhez hasonló beszélgetéseket folytatni. A chatbotok típusai az egyszerűektől az összetettekig terjednek – az utóbbiak már a mesterséges intelligencia széles spektrumát próbálják kiaknázni, és adott esetben a gépi tanulás segítségével a felhasználókkal folytatott beszélgetések alapján képesek lehetnek történeteket is írni. A chatbotok a 2010-es évek kezdetén jelentek meg az üzleti intelligencia egyik specifikus formájaként. A gyakran mesterséges intelligenciával bíró interaktív technológia gyorsan utat talált az online csevegőprogramok világába, és ma már mind több csatornán találkozhatnak a felhasználók vele. A Microsoft például 2014-ben mutatta be Kína legnépszerűbb Weibo nevű közösségimédia-felületén a Xiaolce chatbotját. Rövid idő alatt megfigyelhetővé vált, hogy azok a felhasználók, akik rendszeresen csevegtek a Xiaolcéval személyes témákban, azt írták a rejtett algoritmusoknak, hogy „szeretlek” vagy „hiányzol”. A gépi tanulás segítségével a fejlett chatbot ugyanis meg tudta állapítani a felhasználó üzeneteiből az egyén hangulatát, és

¹¹ Szüts 2016.

¹² Stone 2014.

– amennyiben szükségesnek tartotta az intervenciót – gyakran sikeresen jobb kedvre derítette. A mesterséges intelligenciával bíró chatbotok így – amennyiben arra programozzák őket – képesek lesznek irodalmi szövegek létrehozására, de az ember–gép kommunikáció lejegyzése követheti a nagy írók levelezéseinek sajtó alá rendezését és publikálását is.

Ez a következő eset is alátámasztja: az Eugene Goostman nevű chatbot például 2014-ben elhitette magáról, hogy ember, miközben egy limitált angol szókincsű 13 éves ukrán fiúnak adta ki magát. A londoni Royal Society 30 bírójából ugyanis 10 elhitte, hogy hús-vér személlyel cseveg. A tanulmány címében provokatív módon szereplő Nobel-díjas intelligencia egy fiktív, azonban talán nem oly távoli jövőre utal. Az összetettebb chatbotok képesek bonyolultabb diskurzusokban is részt venni, és úgy programozzák őket, hogy a gépi tanulás révén a korábbi beszélgetéseikből ismereteket szerezzenek, és folyamatosan fejlődjenek. A szofisztikáltabbak szövegértést (natural language processing vagy understanding – N. L. P.) alkalmaznak.

A chatbotokat két csoportba sorolhatjuk. A fa szerkezetű, előre tárolt válaszokat adó (retrieval-based) modellek különböző méretű repozitóriumokat és bizonyos fokú, az emberi memória működését imitáló heurisztikákat használnak a válaszadás során. Ez a heurisztika lehet egyszerű, ekkor alapszintű párosítás szerint érkeznek a válaszok, míg a bonyolult válaszadás során már a gépi tanulás (machine learning) is szerepet játszik. Ezek a rendszerek nem generálnak új szöveget, hanem előre definiált adatbázisból merítenek.¹³ A szövegértésen és szöveggeneráláson alapuló generatív modellek sokkal fejlettebbek, mivel nem előre definiált válaszokon alapulnak, hanem a párbeszédet az alapoktól építik fel. Alapvetően a gépi fordítás (machine translation) technológiáját használják. Az utóbbiak a beszélgetés közben gyakran vétenek

¹³ A natural language processing használata esetében a fejlesztők rendelkezésére állnak például a .NET vagy Java programnyelvek, hogy döntési fa szerkezeteket hozzanak létre, melyek a felhasználókat előre definiált párbeszédútvonalakon vezetik vélhetően egy probléma megoldásáig.

grammatikai hibákat, különösen a hosszú mondatokban. Ezen hibák idővel akkor is feltűnnek a felhasználóknak, ha abból a tényből indulunk ki, hogy az online környezetben létrejött másodlagos írásbeliség jellemzője éppen a nyelvtani szabályok relativizálása.

Irodalom

- Horváth Iván 2000. *Magyarok Babelben*. Szeged: JATE PRESS.
- Licklider, Joseph Carl Robnett 1960. Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics* (1): 4–11.
- Lutz, Theo 1959. Stochastische Texte. *Augenblick* (4): 3–9.
- Stone, Martha L. 2014. *Big Data for Media*. [Report] Institute for the Study of Journalism, <http://reutersinstitute.politics.ox.ac.uk/publication/big-data-media>
- Szombathy Bálint 1977. *A konkrét költészet útjai*. <http://www.artpool.hu/Poetry/konkret/szomitogepes.html>
- Szűts Zoltán – Yoo Jinil 2016. Big Data, az információs társadalom új paradigmája. *Információs Társadalom* (4): 8–28.
- Szűts Zoltán 2013. *A világháló metaforái. Bevezetés az új média művészetébe*. Budapest: Osiris Kiadó.
- Van Mechelen, Vincent 1992. *Computer poetry*. <http://www.trinp.org/Poet/Comp/CompPoe.HTM>

Vidovszky László

DEUS EX MACHINA – a tehetséges számítógép

Ennek az előadásnak a kevésbé hangzatos címe így szólna: kapcsolatok és lehetőségek az informatika és a zene között.

Mindenekelőtt tisztáznunk kell néhány fogalmat. Az informatika esetében viszonylag könnyű a dolgunk: fiatal tudományág jól átlátható kapcsolatrendszerekkel. Elfogadhatjuk tehát, hogy az a tudomány, ami az információ megszerzésével (gyűjtésével), feldolgozásával, tárolásával, sokszorosításával és továbbításával foglalkozik.

A zenével jóval nehezebb a dolgunk. A zene fogalma – megjelenési formájában és tartalmában is – az ember szellemi lehetőségeinek teljes spektrumát átfogja. Egymásnak ellentmondó, gyakran egymást kizáró tulajdonságok egysége. Személyiségünket már a magzati korban alakítja. Egzaktsága csak a matematikáéhoz mérhető, meggyőző ereje, élettani hatása viszont gyakorlatilag mindenkire érvényes. A vele kapcsolatos képzeink, élményeink rendkívül sokfélék. Vannak zenei jelenségek, amik pusztán jelenlétükkel tőlünk függetlenül és általunk nem irányítható módon befolyásolják a viselkedésünket: ilyen például a katonazene vagy a muzak. Ezek általában – már csak feladataik okán is – uniformizálják a gondolkodást. Más esetben a zene szélsőségesen személyes, gondolatisága roppant gazdag – és nem utolsósorban: a liturgia révén a legátszellemültebb gondolatok hordozója és kifejezője. Nyelv, mely önmagát szabadon definiálja.

Természetesen a zenének számos megjelenési formája van, manapság általában a zeneszerző-előadó-hallgató folyamat részeként tekintünk rá. A gyakorlatban ez jellemzően a show-busi-

ness része, rendkívül komplex jelenség, társadalmi hasznossága mindenekelőtt a pénzügyi területen jelentkezik. Ebben a funkciójában már igen messze áll a klasszikusnak mondott zeneművészettől, annak idealizált világtól, jóllehet elsősorban annak eredményeit és esztétikáját használja.

A mai előadásban a zene fogalmát egy jóval egyszerűbb, redukált jelenséggé értelmezem. Zene mindaz, ami hangok kombinációi révén megvalósulhat. Ebben az értelmezésben nincs szó érzelmekről, pátosról, ekstázisról, hanem csak egyszerű frekvenciaviszonyokról, a zene céltalan és univerzális szerepéről. Emiatt a hangzó példákat igyekeztem úgy kiválasztani, hogy azok minél neutrálisabban szóljanak. Ez a hangzás talán fárasztóbb, a hangzás egysíkú és nélküli az előadó személyes tudásának a bőségét, de éppen ezért alkalmasabb egy pusztán demonstráció megszólaltatására.

Közbevetnénk, hogy a zenehallgatás mindenképpen emlékezés is, kultúrája térben és időben is jól körülhatárolható, és nem utolsósorban asszociációs rendszere révén fejt ki a hatását. Nem tudunk zenét hallgatni úgy, hogy kikapcsoljuk korábbi élményeinket, tapasztalatainkat. Nem az adott pillanat, hanem az adott pillanatban megjelenő teljes élet az, ami a zene igazi dimenziója. Ez a zárt összefüggéshalmaz teszi lehetővé, hogy egy kulturális terméket a magunkénak érezzünk, de ez is távolít vagy szeparál el minket más kultúráktól. Úgy tűnik, nincs „elvont” zene.

Mégis akad néhány jelenség, ami el kell hogy gondolkoztathasson. Zubin Mehta egyik interjújában határozottan állítja, hogy az ázsiai ember teljes mértékben megértheti a nyugat-európai zenét, miközben ez fordítva már nem áll. A bizonyítékok sora szinte végtelen: Seiji Ozawától Zubin Mehtáig, Sarah Changtól Yuja Wangig. Mehta nem kevesebbet állít, mint hogy a nyugat-európai zene szerkezete és nyelve oly mértékben egyetemes, minden ember számára érthető és megközelíthető, mint amilyen módon más zenekultúrák nem.

Ezt számos tény látszik igazolni. A hangközök és a számok felfedezése egyaránt Püthagorászhoz („a számok atyjához”) fűző-

dik. A középkor zeneelmélete az univerzum megértésére a hangközök és a számok segítségével törekedett, a polifónia felfedezése pedig többek szerint magának az európai racionális gondolkodásnak volt az alapja. Döntő lépés ezen a területen a notáció felfedezése és kialakulása is. Míg más zenekultúrák a zene lejegyzését, jelekkel való ábrázolását csupán mnemotechnikai segítségként használják, a ma is használatos hangjegyzés – voltaképpen egy stilizált descartes-i koordináta-rendszer – a zenei gondolkodás alapja. A zene vizualizációja tette lehetővé a szabad szólam szerkesztést, ez utóbbi a harmóniai gondolkodás kialakítását.

Ezeket figyelembe véve természetesennek tűnik, hogy a mesterséges intelligencia kutatásainak a közege a zene legyen. Világosan kódolható, relatíve kevés elemszámú eseményből áll. Ezeknek a kutatásoknak az egyik legizgalmasabb kérdése, hogyan tud a számítógép kezelni olyan fogalmakat, amiknek a szimpla meghatározása is vita tárgya, tetszés szerint rokoníthatók, összekapcsolhatók, szabadon újradefiniálhatók, mindemellett nagyon erős érzelmi energiákat is generálnak.

A zene mai megközelítése alapvetően romantikus, annak szélsőséges, egyedi tulajdonságait hagyományozza tovább. Romantikusan itt nem egy zenei gesztusrendszert, még kevésbé zenei stílust értek, hanem a zenével szembeni elvárásokat. A romantikus zene óhatatlanul az egyedire és a kivételesre épít. Nem keresi a zene univerzális összefüggéseit, hanem annak személyre és személyhez szóló lehetőségeit kutatja. Nem véletlen, hogy a 19. század zenéje szinte teljesen nélkülözi a kontrapunktikus gondolkodás eredményeit. Erőtéljes szakítás ez a 17–18. század szemléletével, ahol minden mechanikus elképzelések és szerkezetek révén látszott megvalósíthatónak, mint Vaucanson kacsája, Kempelen beszélőgépe vagy nem utolsósorban Pascal számológépe. Nem véletlen, hogy ezt a gondolkodást a zenére is kivetítették. A különféle automaták mellett elkerülhetetlenül megjelentek a zeneszerző automaták is.

Az algoritmikus zeneszerzésről készült dolgozatok legtöbbször kiemelt helyen említi Mozart *Musikalisches Würfelspiel* cím alatt ismert zenei kockajátékát. A mű valójában egy klasszikus perio-

dizáló szerkezet ütemeinek az egymásutánja, ahol az egyes ütemek zenei anyagai különféle táblázatokból véletlenszerűen kerülnek kiválasztásra.

1. példa [KV 516f]

A hallgató egyformán érez felismerést és zavart. A darab szerkezete, harmóniai bázisa a legáltalánosabb elterjedt kompozíciós gyakorlatot követi, ugyanakkor fordulatai gyakran váratlanok, érdekes, „logikátlanok”. Szokatlan egybeesések vagy különbségek jellemzik ezeket a műveket, ám minél több változatot hallgatunk meg, annál nyilvánvalóbb a gondolkodásmód automatizmusa, a hangzás kiszámíthatósága, a komplexitás és hierarchia hiánya.

Felvetődik a kérdés: vajon mit kerestek a zeneszerzők ezekben a művekben, miféle indítékok munkáltak bennük. Amíg Mozart esetében nyilvánvaló a játék, Johann Philipp Kirnberger művében (*Der allezeit fertige Polonoisen- und Menuettencomponist*) mélyebb összefüggéseket is keres. Kirnberger – Bach koráljainak első közreadója – nemcsak zeneszerző, hanem – talán mindenekelőtt – zeneelméleti szakember is volt. Három hangolási rendszere jelentősen járult hozzá a modern tizenkét fokú hangolás kialakulásához. Műve az algoritmikus komponálás szempontjából sokkal kidolgozottabb, mint Mozarté: nagyobb adatbázisból dolgozik, és a darabok szerkezete is összetettebb.

2. példa [Kirnberger]

Mindkét esetben a kompozíció létrehozója a számítógép „lelke”, a véletlenszám-generátor. Van azonban a történetnek egy másik ága is. A kontrapunktikus szerkesztés különböző formái lehetővé teszik, hogy magán a zenei anyagon végezzünk el bizonyos műveleteket. Idetartoznak a különböző fordítások, projekciók stb. Ezek a transzformációk sok esetben a zenei idő kategóriáit is újraértelmezik. Ebben az esetben a kiindulásul szolgáló zenei anyag már nem valamilyen *objet trouvé*, hanem a szerzői gondolat legalapvetőbb része.

A kontrapunktban rejlő gondolati összefüggések legnagyobb letéteményese Johann Sebastian Bach. A bachi ellenpont tökéletesen megfelel annak a kritériumnak, melyet Beethoven támasztott a zenével szemben: a legmagasabb szintű emberi gondolkodás.

3. példa [BWV 1073]

Ezt a zenét nehéz elképzelni másfajta összefüggésben, mint ami az ő sajátja. Nem lehet feldolgozni valamilyen „cross-over” lemezre, nem lehet „film alá vágni”. Kizárólag a saját maga által megteremtett összefüggésrendszerben értelmezhető.

A kérdés tehát: vannak-e olyan eljárások, algoritmusok, amik segítségével világos, értelmes zenei folyamatok állíthatók elő. A válasz egyértelmű: vannak. Jóval nehezebb a válasz arra a kérdésre, hogy vajon a zeneszerző agyunk fedezi-e föl a zenei törvényszerűségeket, vagy pedig a zene maga alakítja agyunkat a saját törvényszerűségeinek mintájára.

A zenetörténet első számítógép által komponált művét Lejaren Hiller alkotta 1957-ben. A darabot a bemutatón (és azóta is) vonósnégyesen szólaltatták meg azon egyszerű ok miatt, hogy abban az időben a számítógéppel való kommunikáció egyetlen eszköze a lyukkártya volt. A számításokat innen kellett visszafejteni és a zenészek számára olvashatóvá tenni, mivel ebben az időben még nem állt rendelkezésre semmilyen audio interface sem. A vonósnégyes alkalmazását egyébként a darab szerkezete is igazolta: kiinduló anyagnak Hiller ugyanis Bach négyszólamú koráljait használta. Az ezekből leszűrt strukturális következtetések alapján írta meg a programokat.

Az eredményt a zenész szakma szinte egyöntetűen elutasította, de 10 évvel később Hiller részt vett John Cage *HPSCHD* című művének realizálásában, és azóta művének a jelentősége is egyre nyilvánvalóbbá vált.

4. példa: [Hiller: *ILLIAC Suite*]

Néhány szót szólnék saját tapasztalataimról is. Első művem, melyet minden szempontból egy zenei elképzelés digitalizált formájának tarthatok, 1972-ben írtam, egy hangtechnikailag meg lehetőségen bonyolult analóg rendszerre. A címe (*Circus*) a körre utal. A mű egy háromszólamú kompozíció, amelyben egy-egy szólam létrehozásában három-három játékos vesz részt, önállóan változtatva a hang három paraméterét (frekvencia-dinamika-pozíció). Ily módon három (egyre gyengébben érzékelhető) hangzási struktúra jelenik meg egyetlen szólam révén. Az egyes szólamok egymástól teljesen eltérő sebességgel játszanak, emiatt a hangzás végső kialakulása a játékosok számára is kiszámíthatatlan. Az eljárások alapja a léptetés, a számításokhoz programozható számológépet (Texas Instruments) használtam.

Az utóbbi évtizedekben a számítógépes komponálás (CAC) teljesen beépült a zeneszerzői munkámba, de egyfajta sajátos, kevert módon. A zenei anyagok elkészítéséhez rendszerint MAX/MSP programokat készítek, ezeket viszont sok esetben hagyományos vagy más független gesztusokkal vezérlem. A számítógép ebben a munkában több módon is segíthet. Nemcsak az egyszerű szerkezetek továbbfejlesztése, a nagyszámú eseményekből lesűrhető szabályszerűségek kiszámítása, hanem azok dekonstrukciója, más rendszerekbe történő transzportálása révén is.

5. példa: [Vidovszky: *Le piano & ses double V. Orages et arc-en-ciel*]