

Számítógépes irodalomtudomány

Az irodalomtudományi kutatásokban régi ismerős a számítógép, ám mégis újra és újra szükséges bemutatkoznia. A *Helikon* jelen számában a számítógépes irodalomtudomány elmúlt fél évszázadáról és a ma legújabb technológiáiról egyaránt tájékozódhat az olvasó. A kötet tanulmányainak homlokterében a számítógépes filológia elméleti és gyakorlati kérdései állnak. Ha kisebb mértékben is, de a számítógépes irodalomtudomány olyan területei is megjelennek, mint a távoli olvasás, a stilometria vagy az adatbázis-építés.

Julianne Nyhan Tito Orlandival készült interjújának fordítása a digitális bölcsészeti, s egyben a számítógépes filológiai kutatások kanonizált és mitizált ősförását, Roberto Busa kutatásainak tényleges hozadékait vitatja. A beszélgetés a korai nevén számítógépes bölcsészet vagy bölcsészinformatika, ma közkeletű nevén digitális bölcsészet elmúlt közel hét évtizedébe nyújt személyesebb hangvételű betekintést. A számítógépes irodalomtudományról általában, illetve a diszciplínát ért új keletű támadásokról, s azok lehetséges elhárításáról Almási Zsolt ír.

Különszámunk többi cikke a tudományos elektronikus szövegkiadások készítésének kérdéseivel foglalkozik. Maróthy Szilvia bevezető tanulmánya a hagyományos és a számítógépes filológia változásait a vizuálistól az adatszerű megközelítés felé tartó folyamatként mutatja be. Golden Dániel az utóbbi évtizedek eredményeit, beváltott és megvalósításra váró ígéreteit tekinti át, s amellet érvel, hogy a kulturális örökség digitalizációjában korunk humanistáinak, filológusainak különösen nagy szerepe és felelőssége van. Kalcsó Gyula a kéziratos források tudományos elektronikus kiadásának elméleti kérdéseivel foglalkozik, követendő kiadási elveket fogalmazva meg. Fellegi Zsófia és Palkó Gábor az Arany János Emlékév során készült digitális szövegkiadások elkészítésének és webes publikálásának komplex munkafolyamataiba nyújt betekintést, Szénási Zoltán szemléje pedig a hazai tudományos elektronikus szövegkiadások három nagy műhelyének (az ELTE BTK volt Bölcsészinformatikai Önálló Programja, az MTA–DE Klasszikus Irodalmi Textológiai Kutatócsoport, illetve a DigiPhil) kiadványait mutatja be.

A kötet cikkei mind elméleti megfontolásaikkal, mind gyakorlatiasabb javaslataikkal nagyban hozzájárulhatnak a textológiai kutatások újragondolásához, s egyúttal a számítógépes irodalomtudományról szóló diskurzust is megélénkíthetik.

A szám elkészítésében közreműködtek: Tubay Tiziano és Maczelka Csaba.

A kötetet MARÓTHY SZILVIA szerkesztette.

A SZERKESZTŐBIZOTTSÁG

Computational literary studies

In literary research, the computer is an old acquaintance, but it has to be introduced over and over again. In this issue of *Helikon*, readers can learn about the past half century of computational literary studies and the latest technologies in the field. The essays in the volume focus on theoretical and practical issues of computational philology. Other areas of computational literary studies, such as distant reading, stylometry, or database design are also represented to a lesser extent.

The translation of Julianne Nyhan's interview with Tito Orlandi challenges the actual contribution of Roberto Busa's research, which is often referred to as the canonical and mythical source of digital humanities and computational philology. The conversation provides a more personal insight into the past nearly seven decades of the area formerly known as humanities computing or computational humanities, today commonly referred to as digital humanities. Zsolt Almási writes about computational literary studies in general, but also discusses more recent attacks against the discipline, offering strategies to fend them off.

Several articles in our special issue are dedicated to the questions of digital scholarly editions. Szilvia Maróthy's introductory overview presents the changes in traditional and computational philology as a process leading from a visual to a data-driven approach. Dániel Golden reviews the results of the past decades, the already delivered and the yet to be fulfilled promises, and argues that humanists and philologists of our time have a particularly important role and responsibility in the digitization of cultural heritage. Gyula Kalcsó discusses theoretical questions related to the scholarly digital edition of manuscript sources, and defines commendable editorial principles. Zsófia Fellegi and Gábor Palkó provide insight into the complex digital text edition and web publishing workflow carried out during the János Arany Anniversary. Zoltán Szénási reviews the publications of three major Hungarian workshops involved in digital scholarly edition (the former Humanities Computing Independent Program at the Eötvös Loránd University, the Classical Literature Textology Research Group of the Hungarian Academy of Sciences and the DigiPhil).

The articles in this volume, both with their theoretical considerations and with more practical suggestions, can greatly contribute to a revaluation of textual research and at the same time stimulate the discourse on computational literary studies.

The volume was edited by SZILVIA MARÓTHY.

Further contributors: Tiziano Tubay, Csaba Maczelka.

THE EDITORIAL BOARD

Filologia computazionale

Nella ricerca letteraria il computer è una vecchia conoscenza, ma deve essere ancora meglio approfondita. In questo numero di *Helikon*, i lettori possono conoscere gli ultimi cinquant'anni di studi di filologia computazionale e le più recenti tecnologie in questo ambito.

I saggi nel volume focalizzano l'attenzione su questioni teoriche e pratiche di filologia computazionale. Anche se in misura minore, emergono altre aree di studi letterari computazionali come la distant reading, la stilometria o la progettazione di banche dati.

La traduzione dell'intervista di Julianne Nyhan a Tito Orlandi mette in discussione il contributo effettivo della ricerca di Roberto Busa, spesso menzionata come fonte canonica e mitica delle discipline umanistiche digitali e della filologia computazionale. La conversazione offre una visione personale sugli ultimi settanta anni di quell'ambito che precedentemente era conosciuto come informatica umanistica o umanistica computazionale, oggi nota anche come umanistica digitale. Zsolt Almási si è occupato di filologia computazionale in generale, ma anche di confutare i recenti attacchi contro questa disciplina.

Gli altri articoli del nostro numero speciale trattano questioni inerenti alle edizioni critiche digitali. Il saggio introduttivo di Szilvia Maróthy presenta i cambiamenti della filologia tradizionale e computazionale come un processo che procede da un approccio visivo a uno basato sui dati. Dániel Golden esamina i risultati degli ultimi decenni, le promesse già realizzate e quelle ancora da realizzare, e sostiene che gli umanisti e i filologi del nostro tempo hanno un ruolo e una responsabilità particolarmente importanti nella digitalizzazione del patrimonio culturale. Gyula Kalcsó affronta questioni teoriche relative all'edizione critica digitale di fonti manoscritte suggerendo principi editoriali da seguire. Zsófia Fellegi e Gábor Palkó presentano sia i lavori filologici delle edizioni digitali sia il processo di pubblicazione sul web, che si sono svolti nell'anno commemorativo dedicato a János Arany nel 2017. Zoltán Szénási esamina le pubblicazioni delle tre più importanti collaborazioni accademiche ungheresi coinvolte nelle edizioni critiche digitali: il BIÖP dell'Università ELTE, un programma educativo indipendente di informatica umanistica; il MTA-DE Klasszikus Irodalmi Textológiai Kutatócsoport, ovvero il Gruppo di ricerca sulla testologia della letteratura classica dell'Accademia Ungherese delle Scienze e dell'Università di Debrecen; e il progetto DigiPhil.

Gli articoli di questo volume, attraverso considerazioni teoriche e proposte pratiche, possono contribuire notevolmente a una rivalutazione della ricerca testuale e allo stesso tempo stimolare il dibattito sulla filologia computazionale.

Il volume è redatto da SZILVIA MARÓTHY.

Collaboratori: Tiziano Tubay, Csaba Maczelka.

IL COMITATO DI REDAZIONE

TANULMÁNYOK

MARÓTHY SZILVIA

A vizuális és az adatközpontú szemlélet

Bevezető gondolatok

A *Helikon* 1969/3–4. száma ugyan kivételesen nem volt tematikus, ám több tanulmány és recenzió kapcsolódott a „számítógép és a humán tudományok” témaköréhez, ezt emelte ki a címlap is. Közel fél évszázada Petőfi S. János a következő kutatási területekre hívta fel a figyelmet rövid bevezetőjében: automatikus szöveg-előállítás, gépi dokumentáció, számítógépes nyelvelemzés, számítógépes stílus-elemzés.¹ A *Helikon* jelen, *Számítógépes irodalomtudomány* című száma is bizonyítja, hogy ezek a területek a digitális bölcsészet fogalmának és intézményesülésének megjelenésével mit sem változtak.²

A tágabb kontextust egy Tito Orlandival készült interjú fordítása adja, mely többek között a digitális bölcsészeti, s egyben a számítógépes filológiai kutatások kanonizált és mitizált ősforrását, Roberto Busa kutatásainak tényleges hozadékait vitatja. Miben rejlik a valódi paradigmaváltás, milyen volt és milyen ma a bölcsészettudósok viszonya a számítógépes bölcsészettudományokhoz, hogyan változott a számítógépes bölcsészettudományokkal foglalkozó tudományos közösség az évtizedek során? Az interjú a digitális bölcsészeti kutatások elmúlt közel hét évtizedébe nyújt személyesebb hangvételi betekintést. A számítógépes irodalomtudományról általában, illetve a diszciplínát ért új keletű támadásokról, s azok lehetséges elhárításáról Almási Zsolt ír.

A cikkek zöme a számítógépes filológia elméleti és gyakorlati kérdéseit tárgyalja. A téma elsődlegessége jól mutatja, milyen intenzíven foglalkoztatja a történeti források számítógépes feldolgozása a kutatóközösséget. De rávilágít arra is, hogy ezen közösség legfontosabb intézménye, a Textológiai Munkabizottság 2004-ben közzétett kiadási alapelvei³ óta elhanyagolta az eltelt másfél évtizedben rohamosan fejlődő kutatási ágat. Jelen tematikus számban Golden Dániel az utóbbi évtizedek eredményeit, beváltott és megvalósításra váró ígéreteit tekinti át, s amellet érvel, hogy a kulturális örökség digitalizációjában korunk humanis-

¹ PETŐFI S. János, „A számítógépek és a humán tudományok”, *Helikon* 15, 3–4. sz. (1969): 372–373.

² A *Helikon* ezen számáról és a folyóirat más, a számítógépes bölcsészet témájába vágó kiadványairól Golden Dániel emlékezik meg részletesebben a jelen *Helikon*-szám lapjain.

³ „Alapelvek az irodalmi szövegek tudományos kiadásához”, *Irodalomtörténet* 35, 3. sz. (2004): 328–330. A problémáról, illetve az *Alapelvek* alkalmazhatóságáról lásd: MARÓTHY Szilvia, „Tudományos szövegkiadások a hálózaton: Áttekintés”, *Irodalomtörténeti Közlemények* 122, 5. sz. (2018): 617–633.

táinak, filológusainak különösen nagy szerepe és felelőssége van. Kalcsó Gyula a kéziratok források tudományos elektronikus kiadásának elméleti kérdéseivel foglalkozik, követendő kiadási elveket fogalmazva meg. Fellegi Zsófia és Palkó Gábor az Arany János Emlékév során készült digitális szövegkiadások elkészítésének és webes publikálásának komplex munkafolyamataiba nyújt betekintést. Szénási Zoltán szemléje pedig a hazai tudományos elektronikus szövegkiadások három nagy műhelyének (az ELTE BTK volt Bölcsészinformaticai Önálló Programja, illetve az MTA–DE Klasszikus Irodalmi Textológiai Kutatócsoport és a DigiPhil) kiadványait mutatja be.

*

De mihez képest hozott változást a számítógépes rögzítés, feldolgozás és megjelenítés? A '70-es és '80-as években zajló textológiai viták középpontjában az állt, hogy a pozitivizmus örökségét, a kronologikus irodalomtörténeti narratívát a szövegek elgondolásában, értelmezésében, kontextusba ágyazásában és gondozásában felválthatjuk-e, s ha igen, mivel. Az 1976-ban Horváth Iván sajtó alá rendezésében megjelent Balassi-kiadás az eredeti kompozíció, szerzői intenció szerint – illetve azt platonikus értelemben „eszményítve” – közli a verseket. Mint utószavában írja:

Azért nevezhető kísérletinek ez a kiadás, mert – a korábbi kiadói gyakorlatól eltérően – Balassi Bálint verseit nem a soha pontosan meg nem állapított kronológia alapján, hanem elvszerűen a költő saját versgyűjteményeinek (ciklusainak) sorrendjében közli.⁴

A kiadás a szemléletváltással nem csupán azt kezdeményezi, hogy a versek időbeli elhelyezése helyett azok költői dispositiójára irányuljon a figyelem, hanem (Bori Imre szavaival) a költő „művészi öntörvényeinek leírása kerül előtérbe az eddigi életrajz-központú szemlélet ellenében.”⁵ Szigeti Csaba az *Appendix Balassianá*-ban a kronologikus gondolkodás problematikusságát a régiség irodalmát illetően Koháry István és Beniczky Péter költészetén mutatja be. Alább elméleti megfonto-

⁴ HORVÁTH IVÁN, szerk., *Balassi Bálint összes versei: a versek helyreállított eredeti sorrendjében* (Novi Sad: Újvidéki Egyetem, 1976), 143. Az eszményített kompozíció egyúttal a Gerézdi Rabán és Klaniczay Tibor 3×33-as és Varjas Béla 2×33-as szerkezeti hipotézisére adott válasz.

⁵ Uo., 145. Bori Imre a kiadáshoz írt előszóban hozzászeli, hogy szándékuk nem magában álló, ez az igény az 1976-ban Egerben megrendezett Balassi-értekezleten is megfogalmazódott. Vö. *Irodalomtörténeti Közlemények* 80, 5–6. sz. (1976). Az *Itk* ezen számában Szörényi László Balassi török bejтей mint kompozíciót elemzi. Tanulmányának keletkezéséről ezt írja: „Ez a dolgozat a Balassi Bálint életművét tárgyaló tudományos ülésszakon, Egerben, 1976. május 14-én elhangzott előadás kibővített változata. Létrejöttében szerepet játszott Horváth Iván baráti készítése, aki e sorok szerzőjével együtt már évek óta törekszik arra, hogy a klasszikus költők életművének kiadásában jelenleg kizárólagos időrendi elv mellett a kellő súllyal érvényesüljenek a kötetkompozíció követelményei.” SZÖRÉNYI László, „»Valahány török bejt«”, *Irodalomtörténeti Közlemények* 80, 5–6. sz. (1976): 706–713, 706.

lásaiból ragadok ki szemelvényeket a kronologikus és a kompozíciós értelmezés egymásnak feszülését illusztrálандó.

[... hogy] a kronologikus sor szemléletszerűen mennyire áthatja az irodalomtörténetet, úgy gondolom, bizonyítás nélkül is megálló közhely és igazság.

[...] az adatvirágok elvileg végtelen mennyiségben fűzhetők fel a relatív kronológia madzagjára, ami a filológus számára a gazdagság (a kimeríthetlenség) érzetének alapját alkotja meg. Az adat vagy tény, esetleg szöveg beillesztése az időrendi sorba: a behelyezés. E művelet steril irodalomtörténeti eljárásnak tűnik, holott – könnyen belátható, hogy – ideologikus, manipulatív gesztus, sőt minden későbbi értelmezés bázisa. [...] A kronologikus gondolkodású irodalomtörténetész lényegében egyetlen pillanat érdekl, egyetlen pillanatot értelmez: a mű keletkezésének pillanatát.⁶

Noha a kronologikus gondolkodást tekinthetjük egyeduralkodónak az ezredforduló környékéig,⁷ fontos látni, hogy a textológiai kritikában korábban is fel-felmerült ennek tarthatatlansága. Szigeti Csaba idézett cikkében Toldy Ferenc egy közleményét hozza fel példaként, melyben felhívja a figyelmet Beniczky addig ismeretlen korai költeményeire, s a *Magyar rithmusok* verseinek feltehetőleg jóval korábbi keletkezésére. Ez alapján módosítja a Beniczkyról, a kötet korabeli költészetébe ágyazottságáról alkotott képet: „Beniczkyt tehát ezen túl az irodalmi történetirő nem Zrínyi után, hanem elibe fogja helyezni, s Rimai mellé, kinek ifjabb kortársa volt.”⁸

Horváth Iván fent hivatkozott, Balassi kötetkompozíciójával foglalkozó cikkében Klaniczay Tibor 1957-es tanulmányát idézi, melyben Eckhardt kiadása apropóján kitér Balassi és Rimay nagykompozíciós törekvéseire is. Mint írja, Balassi és filológiája ebben is úttörő: „A kronologizmusnak Balassi vetett véget. Klaniczay Tibor (1957) javaslatot tett a költő »maga kezével írt könyvének« verssorrendjére.” Klaniczay így foglalja össze hozzászólásának célját: „felülvizsgálom Balassi- és Rimay kéziratos és nyomtatott szöveg-hagyományának egész problematikáját, s ezen az úton megpróbálom a kronológia és szövegkritika terén eddig elért ered-

⁶ SZIGETI Csaba, „Appendix Balassiana: Kronológia, tradíció, hagyománytudat a XVII. századi Balassi-követő nemesi költészetben”, *Irodalomtörténeti Közlemények* 89, 6. sz. (1985): 675–687, 675–676.

⁷ A narratológiai-textológiai fordulatról lásd legújabban Horváth Iván cikkének bevezetőjét, melyben nem annyira a pozitívizmus örökségével való leszámolásként, inkább a megszállás időszakának ideológiai meghatározottságától való elszakadásként exponálja a bekövetkezett változást. HORVÁTH Iván, „Balassi Bálint verseinek fragmentumi”, *Credo* 24, 3–4. sz. (2018): 18–39.

⁸ TOLDY Ferenc, „Beniczky Péter kora, s némely fontos kicsiség”, *Új Magyar Múzeum* 3, 1. sz. (1853): 352–356, 355. Igaz, ahhoz, hogy Rimay kortársának tekintsük, el kellene fogadnunk Toldy azon spekulációját („De ha okoskodni szabad – pedig szabad a historiában is, adatok létében és nem létében! ...”), hogy Beniczky 1632-re elhunyt. Ezt azonban csupán egy possessorbejegyzésre alapozza, s pár évtizeddel később Komáromy közlése nyomán kiderült, Beniczky 1664-ben hunyt el. KOMÁROMY András, „Adalékok Beniczky Péter életéhez”, *Történelmi Társulat* 11, 3. sz. (1888): 435–449, 448.

ményeket továbbfejleszteni.”⁹ Mind Toldy, mind Klaniczay esetében láthatjuk, hogy a szándék nem forradalmi, nem törekszik kizárólagosságra: a kronológiai és a kompozíciós értelmezői narratívát kívánja ötvözni.

A számítógépes szövegfeldolgozás lehetőségével nemcsak a tudományos szövegkiadás fogalma kérdőjeleződött meg, de a szövege is. A médiumváltással összefüggésben jelenik meg a szöveg pluralitásának felismerése,¹⁰ vagy legalábbis láthatóvá tétele, a szöveg határainak megkérdőjelezése és a hipertext-jelenség, valamint a szöveg mint hierarchikus felépítmény modellezhetősége is.¹¹ Bernard Cerquiglioni szövegvariánsoknak szentelt könyvében már kijelöli azokat a lehetőségeket, melyeket a számítógépes feldolgozás a könyv megjelenésekor (1989) ígért a középkor filológiája számára¹² – például a szövegvariánsok párhuzamos megjelenítését, az annotálás határtalanságát, indexek és konkordanciák generalitását. Karina van Dalen-Oskam mintegy negyedszázaddal később tanulmányában azt vizsgálja, milyen módon és mértékben valósultak meg ezek:¹³ a párhuzamos megjelenítésre ma számos példát találhatunk, a különféle annotációknál pedig legfeljebb az okozhat gondot, hogyan helyezték el azokat, milyen interaktív lehetőségeket biztosítsanak a felhasználó számára, hogy a befogadás gördülékenyen menjen.

⁹ KLANICZAY Tibor, „Hozzászólás Balassi és Rimay verseinek kritikai kiadásához”, *MTA Nyelv- és Irodalomtudományok Osztály Közleményei* 17, 1–4. sz. (1957): 265–338, 266.

¹⁰ A plurális szövegfogalomhoz lásd: Bernard CERQUIGLIONI, *In the Praise of the Variant*, trans. Betsy WING (Baltimore, London: John Hopkins University Press, 1999). Eredeti kiadvány: Bernard CERQUIGLIONI, *Éloge de la variante: Histoire critique de la philologie* (Paris: Ed. du Seuil, 1989); HORVÁTH Iván, „Szöveg”, 2000, 11. sz. (1994): 42–53; HORVÁTH Iván, „A hálózati kultúra fenomenológiája: Egy általános textológia vázlata”, *Gépeskönyv*, 2001, <http://magyar-irodalom.elte.hu/vita/thi2.html>. Ebben a szellemben készült továbbá a régi magyar versek repertóriumára is: *Répertoire de La Poésie Hongroise Ancienne I–II*, dir. Iván HORVÁTH (Paris: Le Nouvel Objet, 1992). Hálózati kiadás: ua., 4.0 v, hozzáférés: 2020.03.21, <http://rpha.oszk.hu/>.

¹¹ Steven J. DE ROSE, David G. DURAND, Elli MYLONAS and Allen H. RENEAR, „What is Text, Really?”, *Journal of Computing in Higher Education* 1, no. 3. (1990): 3–26. Az OHCO-modellhez (Ordered Hierarchy of Content Objects) és verzióhoz lásd még: Allen RENEAR, Elli MYLONAS and David DURAND, „Refining Our Notion of What Text Really Is”, 1993, <http://cds.library.brown.edu/resources/stg/monographs/ohco.html#sec13>. Filológiai történeti áttekintés: LABÁDI Gergely, „A filológiai tudás formái”, in *Textológia, filológia, értelmezés: Klasszikus magyar irodalom*, szerk. CZIFRA Mariann és SZILÁGYI Márton, Csokonai Könyvtár: Bibliotheca Studiorum Litterarium 55, 173–190 (Debrecen: Debreceni Egyetemi Kiadó, 2014).

¹² Illetve részint be is váltott addigra, lásd pl. a Roberto Busa által az 1950-es években kezdeményezett *Corpus Thomisticumot*, vagy az 1970-es években indult már idézett *Répertoire de La Poésie Hongroise Ancienne*-t. Roberto BUSA, „L’Index Thomisticus e l’informatica filosofica”, *Revue Internationale de Philosophie* 27, no. 1. (1973): 31–36; RPHA MUNKACSOPORT, „Szegedi kísérlet a XVI. századi magyar vers gépi feldolgozására”, *Irodalomtörténeti Közlemények* 84, 5–6. sz. (1980): 630–638.

¹³ Karina VAN DALEN-OSKAM, „In Praise of the Variant Analysis Tool: A Computational Approach to Medieval Literature”, in *Texts, Transmissions, Receptions: Modern Approaches to Narratives*, eds. André LARDINOIS, Sophie LEVIE, Hans HOEKEN and Christoph LÜTHY, *Radbound Studies in Humanities* 1, 35–54 (Leiden: Brill, 2015), <https://www.jstor.org/stable/10.1163/j.ctt1w76wgh.7>. A jelen folyóirat számban hasonlóra tesz kísérletet Szénási Zoltán cikke.

A visszakereshetőség az, ami talán a leginkább felülmúlta a korabeli várakozásokat, hiszen nemcsak tetszőleges indexek generálhatók a rögzített adatokból, hanem a rendelkezésre álló lekérdező eszközök és a ráépülő szoftveres támogatás (például adatvizualizációk) is olyan rálátást adnak a szövegekre és azok adataira, melyek új, a korábbi eszköztárral nem hozzáférhető felismeréseket hozhatnak. Jelen pillanatban azonban a filológia még mindig időigényes és csak részben automatizálható műveletek sorából áll, így a nagy méretű (*big data* szintű), filológiai-lag is megbízható szövegkorpuszokért, s az azokon végezhető vizsgálatokért még dolgozni kell.

Az elmúlt közel huszonöt év szövegkiadásaiból válogatva¹⁴ úgy tűnik, többségben vannak azok, amelyek nem kezelik külön, tehát nem regisztrálják és jelenítik meg az eredeti forrásbeli szövegsorrendet. Ehelyett jellemzően tematikus, szerkezeti és formai szempontok szerint indexált, illetve kereshető megjelenítéssel találkozunk, különösen a 2010-es évek előtti, HTML-alapú kiadások esetében. A szövegek egymásutánosságára a fájlközpontú (számítógépes adatkezelési egységekből építkező) szemlélet miatt, mely szerint egy szövegegység egy fájl, ritkán találunk reflexiót. Ez azt jelenti, hogy az egyes szövegforrásokat feldolgozó fájlok nem tartalmaznak információt a szövegforrásban (kéziratban vagy nyomtatványban) előttük lévő, illetve utánuk következő szövegről, esetleg magáról a szövegforrásról sem. Többnyire ezen adatok is eljutnak más módon (például index oldalon, kísérőtanulmányban) a felhasználóhoz, de a szöveg kódolásában ezek az adatok nincsenek explicite jelen.

Az a kérdéskör tehát, hogy milyen módon adjunk közre szépirodalmi szövegeket, a kronológia és így közvetetten az életrajzi referencialitás, vagy a szerzői intenciók elsődleges figyelembevételével, megszűnt problémának látszani. Talán a szöveg plurális megközelítése vonta el erről a figyelmet. A számítógépes textológia egy forrásközpontú, pontosabban a források kevésbé hierarchikus felfogású közreadási tevékenységére váltott át. A számítógépes feldolgozás, illetve a webes sajtó alá rendezés és megjelenítés eltérő jellege miatt a szövegek feletti kompozíció explicit regisztrálása gyakran csak a kiadásban kódolt szövegeken kívül kap helyet – a művek tulajdonképpen darabokra esnek, minden drámaiság nélkül.

A helyzet dramatizálására azért sincsen okunk, mert ez a hiányosság nem a médium sajátossága. Ha jobban meggondoljuk, explicite a nyomtatott kiadásokban sincs általában rögzítve a szövegek egymásutánosságának ténye (tehát az egymásutánosság mint a szövegegységek egy tulajdonsága) – egészen addig, míg az említett fordulatok, a textológia kritika a '70–'80-as években azzá nem tette a

¹⁴ A két legátfogóbb katalógus digitális szövegkiadásokra: Patrick SAHLE, „A catalog of Digital Scholarly Editions, v 3.0” (2008–), <http://www.digitale-edition.de/>; Greta FRANZINI, „Catalogue of Digital Editions” (GitHub, 2012–), https://github.com/gfranzini/digEds_cat, ennek böngészhető változata: <https://dig-ed-cat.acdh.oew.ac.at/>. Cikkem ezen részében a 2019-ben vitára bocsátott disszertációm eredményeire támaszkodom. MARÓTHY Szilvia, *Szerzői verskötetek a 17. századi magyar irodalomban* (ELTE BTK Régi Magyar Irodalom Tanszék, 2019), 139–152, <https://edit.elte.hu/xmlui/handle/10831/44817>.

korábban inkább implicit létmódú állításokat (ti. hogy a költői életmű elemeinek sorrendje kronologikus). A kronologikus irodalomtörténeti gondolkodás hatással volt a szövegkiadásokra, de fordítva is: a kronologikus szövegkiadások hatással voltak (vannak) az irodalomtörténeti gondolkodásra. A számítógépes filológia adatközpontú szemlélete az, ami felhívta a figyelmet a hagyományos filológia vizualitáshoz, tipográfiai megoldásokhoz való erős kötődésére. A papíralapú kritikai kiadás magától értetődően alkalmaz tipográfiai megoldásokat filológiai megállapítások kifejezésére.¹⁵ Így ezek a megállapítások adatszerűen nincsenek rögzítve a kiadásban, vagy nem a szöveghez kapcsolódnak közvetlenül. Ezt a funkciót is többnyire a kísérőtanulmány látja el.

A kísérőtanulmány számos eleme a számítógépes szövegtárolás esetében annak metaadatai közé vándorol. A TEI Header, mely a kódolt szöveghez kapcsolódó metaadatokat rögzíti,¹⁶ épp annak érdekében jött létre és finomodott az évek során, hogy a tanulmányban összegzett információkat szegmentálja, s ezáltal újrafeldolgozhatóvá, összehasonlíthatóvá és nem utolsósorban a weben könnyebben megtalálhatóvá tegye. Mindez természetesen nem azt jelenti, hogy a kísérőtanulmányt egy metaadathalmaz váltja fel (egyelőre), hanem azt, hogy a számítógépes feldolgozás megköveteli a filológiai eljárásokra való reflektálást strukturált adathalmaz formájában is.

A hagyományos és a számítógépes filológia összevetésére visszatérve, a tudományos kiadások elemzése arra világít rá, hogy a hagyományos filológia vizualitásból kiinduló, vizuális eszközökre alapozó szemlélete a számítógépes eszközökkel létrehozott tudományos kiadásokra is sok esetben jellemző. Bár a kiadások és a filológiai gyakorlatok igen eltérőek, talán mégis ábrázolhatjuk ezt egy folyamatként, mely a vizuálistól az adatszerű reprezentáció felé halad. A folyamat egy jelentősebb állomása a HTML kódolásról az XML kódolásra való áttérés, mely által különvált a tudományos adat és annak megjelenítése. Ezzel párhuzamosan egy jelentős ellenmozgás, visszalépés a HTML-ről a PDF formátumokra való áttérés, mely újra a vizuális kódokat részesíti előnyben. Előrefelé egy következő állomás az XML szintaxis meghaladására törekvő ún. stand-off markup, illetve a Linked Open Data, vagy gráfalapú szemlélet, mely még inkább a szöveg sokféleségének (multiplicitásának és pluralitásának) megragadására koncentrál.¹⁷

Míg a HTML kódolással az alapvető probléma az volt, hogy a tartalom és annak megjelenítése nem válik külön, az XML-lel az, hogy nem kezeli jól az átfedé-

¹⁵ A hagyományos és a digitális filológia szemléleti ütközéseire hoz szemléletes példákat: Magdalena TURSKA, James CUMMINGS and Sebastian RAHTZ, „Challenging the Myth of Presentation in Digital Editions”, *Journal of the Text Encoding Initiative* 9 (2016), <https://doi.org/10.4000/jtei.1453>.

¹⁶ „The TEI Header”, in TEI CONSORTIUM, *P5: Guidelines for Electronic Text Encoding and Interchange*, v 4.0.0, hozzáférés: 2020.03.19, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/HD.html>.

¹⁷ „Stand-off markup”, TEIWiki, hozzáférés: 2020.03.21, https://wiki.tei-c.org/index.php?title=Stand-off_markup.

seket, s a beágyazottság miatt redundánssá, zavarossá válhat.¹⁸ Az XML szintaxis számára megalkotott TEI kódolás pedig úgy látszik, nem egészen váltotta be az olyan vele kapcsolatos reményeket, mint a platformfüggetlenség, az újrafelhasználhatóság, vagy az interoperabilitásra való alkalmasság. Ezért került az utóbbi években a figyelem középpontjába a szövegek közötti (*inline*) helyett a szövegen kívüli (*stand-off*) kódolás, markup. A módszer lényege, hogy külön rétegekként és külön számítógépes fájlkként is kezeli a nyers szöveget (*plain text*) és a hozzá tartozó szerkezeti, tartalmi, szövegkritikai stb. állításokat, így lehetővé válik a szöveggel kapcsolatos, egymással a szöveg linearitása miatt átfedésben lévő állítások megfogalmazása. A *stand-off* maga nem új keletű, az 1960-as évek végéhez és Ted Nelson munkásságához kötik első megjelenését,¹⁹ a humán tudományok területén azonban elsősorban a nyelvészek alkalmazták eddig.²⁰

Az alábbiakban két nem kifejezetten nyelvészeti alkalmazást mutatok be röviden, Desmond Schmidt, valamint Elena Spadini és Magdalena Turska megoldási javaslatát. Schmidt a kódolásnak négy szintjét különíti el: metaadat, szöveg (variánsok), jegyzetelés (értelmezői réteg), szövegtulajdonságok (például bekezdések, kiemelések).²¹ A TEI XML kódolás szerint ezek egy fájlban belül találhatóak, az adott szövegforrás(ok)hoz kapcsolódnak. A *stand-off* módszer szerint azonban szétválnának különálló, más formátumú fájlokra: metaadat, nyers szöveg, különféle annotációk (például értelmező jegyzetek, szövegtulajdonságok, szövegkritikai apparátus).

Kérdés azonban, ha a szövegkritikai apparátust leválasztjuk a szövegről (szövegforrásunk átiratáról), mi marad, mi az a nyers szöveg, melyhez az annotációkat

¹⁸ Az XML szintaxisához lásd: <https://www.w3.org/XML/>. Példa átfedésre:

Az ifju visszaretent: „Hah! megint te?

Mindig te!” és elfordítá szemét, (*Bolond Istók*, I/44/1–2, AJÖM III/146.)

A két verssorban egy idézet szerepel, mely az egyik versorból a másikba nyúlik át. Ez az átfedés az XML szintaxis szerint nem lehetséges, az egyes elemek egymásba kell hogy ágyazódjanak, itt azonban nem ez a helyzet. Hibás átírás lenne tehát:

<|>Az ifju visszaretent: <q>Hah! megint te?</|><|>Mindig te! </q>és elfordítá szemét,</|>

Ezt át lehet ugyan hidalni többféleképpen, ám egyik sem nyújt természetes, illetve az XML feldolgozás szempontjából hatékony megoldást. Például üres <q> elemek beszúrásával, melyek kezdő (sID) és záró (eID) attribútumot kapnak.

Pl. <|>Az ifju visszaretent: <q sID="idezet1" />Hah! megint te?</|><|>Mindig te!<q eID="idezet1" /> és elfordítá szemét,</|>

¹⁹ Desmond Allan SCHMIDT, „Using standoff properties for marking-up historical documents in the humanities”, *Information Technology* 58, no. 2. (2016): 63–69, 64, <https://doi.org/10.1515/itit-2015-0030>.

²⁰ A megoldás újszerűségét mutatja, hogy a feltehetőleg első nagyobb volumenű TEI XML–standoff konverzióról egy tavalyi cikk számol be. Giuseppe G. A. CELANO, „Standoff Annotation for the Ancient Greek and Latin Dependency Treebank”, in *Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage – DATeCH2019*, 149–153 (Brussels: ACM Press, 2019), <https://doi.org/10.1145/3322905.3322919>.

²¹ Desmond Allan SCHMIDT, „Standoff properties as an alternative to XML for digital historical editions” (*Semantic Scholar*, 2016), 1–23, 6, <https://www.semanticscholar.org/paper/Standoff-properties-as-an-alternative-to-XML-for-Schmidt/6056afc3c25fc0e9b3e677c04ea4bc34b8151ab>.

kapcsoljuk? Ez ugyanis elvben nem tartalmazhat semmilyen formázást, szegmentálást, de szövegváltozatokat sem. Ha egy szövegforrás javításokat, törlést, betoldást tartalmaz, az egyes szövegállapotokat miként rögzítjük stand-off modellben? Schmidt javaslata szerint ilyenkor egy-egy szövegforrásnak több átíratát kell elkészítenünk, melyek megfelelnek az egyes szövegállapotoknak.²² Ez az elképzelés azonban azt feltételezi, hogy az adott szövegegeszen esett változtatásokat rétegekké tudjuk szétfésülni. Például a szöveg eleji törlésről és a szöveg végén található betoldásról tudjuk, hogy ugyanazon szövegréteghez tartoznak, tehát közel egy időben keletkeztek – erről ritkán vannak információink. A nyers szöveg Schmidt megoldásában ráadásul – mint azt maga is elismeri – nem teljesen mentes a szövegtulajdonságoktól, metaadatoktól, a paragrafushatár jelölésére például a Markdownhoz hasonlóan a dupla sorközt vagy nagyobb térköz alkalmazását javasolja, illetve az oldalszám nyers szövegben való szerepeltetését bizonyos esetekben indokoltan tartja.

Bár elméleti szempontból nagyon vonzó a stand-off markup azon elképzelése, hogy a szöveget és a rá vonatkozó állításainkat elkülönítse, ennek textológiai és technológiai gyakorlatba ültetése egyaránt rengeteg kérdést, problémát vet fel. A stand-off ugyanúgy magában rejtje a kód sokszor felesleges bonyolításának lehetőségét. Szemléletes példákat hoz a lekérdezés bonyolultságára Spadini és Turska: például míg egy adott versszöveg negyedik sorának XPath lekérdezése TEI XML-ben 13 karakter hosszú, stand-off rendszerben háromsornyi, tehát 2–300 karakternyi is lehet.²³ A szerzők azonban elismerik, hogy a TEI XML is gyakran válik áttekinthetetlené vagy redundánssá. Elismerve a stand-off előnyeit a két technológia ötvözésére tesznek javaslatot. A TEI XML jelenlegi két fő részét (TEI Header és Text) egy harmadikkal bővítenék ki, melybe a szövegen kívüli annotációk kerülnének, rétegekként rögzítve. A <teiHeader> továbbra is a metaadatokat tartalmazza, a <text> a (mondatra, sorra, szóra stb.) szegmentált szöveget, ahol minden szegmens egyedi azonosítót kap. A harmadik rész, a <standoff> tartalmazná rétegekre bontva az egyes szövegszegmensekre, intervallumokra vonatkozó annotációkat.²⁴ Például az egyik réteg a szöveg szerkezetét írja le (cím, strófa, verssor stb.), a másik a szövegtulajdonságokat (idézet, kiemelés stb.), a harmadik a szövegkritikai megállapításokat (törlés, betoldás, hiány stb.).

Schmidt, illetve Spadini és Turska megoldásait összevetve az egyik szembetűnő különbség, hogy míg előbbi a szövegek többszörözésével éri el, hogy a szövegalakulás egyes állomásait rögzítse, addig utóbbi – a TEI hagyományait követve –

²² Uo., 7–11.

²³ Elena SPADINI and Magdalena TURSKA, „XML-TEI Stand-off Markup: One Step Beyond”, *Digital Philology: A Journal of Medieval Cultures* 8, no. 2. (2019): 225–239, 234, <https://doi.org/10.1353/dph.2019.0025>.

²⁴ Uo., 228–230. Az elemkészletet bevezető cikk: Javier POSE, Patrice LOPEZ and Laurent ROMARY, „A Generic Formalism for Encoding Stand-off Annotations in TEI”, 2014, <https://hal.inria.fr/hal-01061548/>.

az annotációban rögzíti azokat. Mindkét esetben kérdéses marad, mit tekinthetünk alap- vagy nyers szövegnek, illetve hogyan körvonalazzuk az egyes szövegrétegeket. A másik fontos különbség a szövegnek mint adatok halmazának hierarchikus vagy nem-hierarchikus értelmezése. A TEI XML és a mögötte rejlő OHCO-modell hierarchikus felépítményként értelmezi a szöveget. A stand-off markup nem, vagy csak részben hierarchikusként, elismerve, hogy a szöveg csak külön értelmezési keretek rétegeiként írható le. Mindkét modell helytálló lehet: a különféle textológiai hagyományok és a szöveg egyéni vonásai között kell egyensúlyt teremteni, s ahhoz érdemes technológiát választani.

Lehet ugyan, hogy az XML-kódolásnál a stand-off technikailag jobban támogatja az együttműködést és a platformfüggetlenséget, de az is előfordulhat, hogy a technikai mellett az emberi tényezőkben kell keresnünk a hibát. A tudományos elektronikus kiadást készítőik ugyanis csak ritkább esetben teszik közzé az eredeti, kódolt szövegállományt és járulékos fájljait (például TEI specifikációt),²⁵ amelyre a webes megjelenítés épül. Kutatásuk leginkább innovatív, s a tudományos közösség számára jelentős eredményeik rejtve maradnak, nem vizsgálhatók. Mintha egy nyomtatott kiadásban csak a jegyzetek felét tennék közzé, vagy az alkalmazott szöveggéközlési elveknek csak egy részébe avatnák be az olvasót. Az összehasonlítás végett: Franzini jelenleg 304 tételes katalógusának csak valamivel több mint 20 százaléka (63 darab) biztosítja a TEI XML források letöltését. Ebben a katalógusban természetesen számos HTML alapú és más, nem TEI XML szabványt alkalmazó hálózati kiadás van (nem TEI XML összesen 135). Ha a TEI XML-ben készült kiadásokat nézzük, az arány már biztatóbb, azok több mint 37 százaléka szabadon hozzáférhető (63/169). Sajnos egy magyar sincs köztük.²⁶

A számítógépes feldolgozás, elemzés és webes közzététel egy másik (nem csupán a számítógépes irodalomtudományt érintő) megoldásra váró problémája annak kutatásértékelésben elfoglalt helye. A tudományometriai és kutatásértékelési szempontokat magába olvasztó rendszer, a Magyar Tudományos Művek Tára (MTMT) a számítógépes eszközökkel létrehozott kutatási eredményeket nem, vagy alig díjazza, függetlenül azok tartalmától, minőségétől és publikáltságától. Nemcsak a kutatás viseli tehát magán a vizualitás, a vizuális alapú gondolkodás nyomait, hanem a kutatásértékelés is. A grafikus megjelenített tartalom (főleg, ha nyomtatott könyv): tudományos közlemény. A digitális fájlok, adatok halmaza (ha nem nyomtatott – s ez a jellemző): egyéb. A döntés során a tartalmi jellemzőket a formaiak írják felül. A közlemény és a közleményekre történő hivatkozások elszámolása tekintetében egyaránt hátrányban van még ma is az a magyarországi kutató, aki korszerű technikai eszközök bevonásával végzi és teszi közzé kutatásait.

²⁵ Utóbbira jó példa a *Deutsches Textarchiv* példásan részletes, mindenki számára hozzáférhető dokumentációja. Martin GRÖTSCHEL, Wolfgang KLEIN und Alexander GEYKEN, *Deutsches Textarchiv (DTA)* (Berlin: Brandenburgische Akademie der Wissenschaften, 2007–), <http://www.deutschestextarchiv.de/>.

²⁶ Az adatok Franzini hivatkozott katalógusából származnak, hozzáférés: 2019.05.03.

Egy konkrét példát említve: az *Ómagyar Mária-siralom* TEI XML-kiadását²⁷ többéves kutatói munka előzte meg. A kiadás egy OTKA-projekt (119355) támogatásával valósult meg, megjelenése előtt szakmai fórumokon vitára bocsátottuk eredményeinket,²⁸ a kiadás bírálói, szakmai tanácsadói a téma kiemelkedő szakértői voltak. A kiadás megfelel az MTA Textológiai Munkabizottsága által a kritikai kiadással szemben támasztott követelményeknek (például: bevezető tanulmány, mely ismerteti az előzményeket és a szövegekölési elveket, kritikai apparátus). Végezetül – számos internetes szövegkiadással ellentétben – DOI azonosítóval is rendelkezik. Mindezek ellenére nem számít tudományos közleménynek, az MTMT adminisztráció szerint azért, mert nincs publikálva, noha ez az egyetlen jelenleg szabadon hozzáférhető, a weben közzétett hazai TEI XML szövegekölés.²⁹

Változást minden bizonnyal valóban a 2020-as évek hoznak, erre utal, hogy a tudomány nyílt hozzáférést támogató Európai Unió ajánlások idén a magyarországi pályázati rendszert, s így a kutatási infrastruktúra felelős intézményeit is elérték. A kutatási adatkezelés (*research data management*), s ennek jelenleg legkidolgozottabb ajánlása, a FAIR alapelvek (Findable, Accessible, Interoperable, Reusable)³⁰ remélhetőleg rövidesen általános gyakorlat részeivé válnak.³¹

A számítógépes filológia megoldatlan kérdéseire a jelen *Számítógépes irodalomtudomány* című szám több cikke is felhívja a figyelmet. Alapvető elméleti problémákra térnek vissza újra és újra, melyek a nemzetközi szakirodalomban is tematizálódtak az elmúlt közel fél évszázadban. A hazai textológiai szakirodalom az utóbbi két évtizedben³² viszonylag kevésbé foglalkozott a számítógépes filológia elméletével ahhoz képest, hogy a webes szövegkiadások mennyisége – s lassan tudományos értéke is – jelentősen növekedni kezdett. Az 1990-es évektől a 2000-es évek közepéig tartó első fellendülést, mely elsősorban a Horváth Iván vezette Bölcsészinformatikai Önálló Program kiadványaihoz és rendezvényeihez kötődik,

²⁷ HORVÁTH Iván és MARÓTHY Szilvia, szerk., *Ómagyar Mária-siralom: kritikai kiadás*, v1.0 (GitHub-Zenodo, 2018), <http://doi.org/10.5281/zenodo.1287583>. Ez a kiadás, s más, újabban keletkezett kiadványok Franzini már hivatkozott katalógusába még nem kerültek fel.

²⁸ Pl. HORVÁTH Iván és MARÓTHY Szilvia, „Felkészülés a HBK kiadására: az ÓMS próbakiadása”, *A Pray-kódex-munkacsoport első műhelykonferenciája*, Budapest, Országos Széchényi Könyvtár, 2017. június 13; MARÓTHY Szilvia, „Az ÓMS elektronikus kritikai kiadása”, *III. Forráskutatás, forráskiadás, tudománytörténet konferencia*, Budapest, Eötvös Loránd Tudományegyetem, 2017. november 16–17.

²⁹ Ti. semelyik másik tudományos szövegkiadást TEI XML-ben kódoló, weben közzétevő személy, illetve intézmény nem teszi hozzáférhetővé ezen kiadások lényegét, a TEI XML forrásfájlokot, lásd fentebb.

³⁰ Mark D. WILKINSON, Michel DUMONTIER, IJsbrand Jan AALBERSBERG, Gabrielle APPLETON, Myles AXTON, Barend MONS et al., „The FAIR Guiding Principles for Scientific Data Management and Stewardship”, *Scientific Data* 3 (2016): 160018, <https://doi.org/10.1038/sdata.2016.18>.

³¹ A kutatási adatkezelés, nyílt hozzáférés, FAIR alapelvek témaköréhez lásd: MARÓTHY Szilvia, „A nyílt és a zárt tudományról”, in *Kulturális iparágak, kánonok és filterbuborékok*, szerk. BÁRÁNY Tibor et al. (Budapest: BME-ELTE, 2020, megjelenés alatt).

³² A témával foglalkozó cikkeket lásd a jelen kötet válogatott bibliográfiájában. A hazai tudományos elektronikus kiadások bibliográfiája elérhető: <https://www.zotero.org/groups/2199751/> és https://github.com/marothyszilvi/digEds_hun. Vö. MARÓTHY, „Tudományos szövegkiadások a hálózaton...”.

úgy tűnik, csak most, a 2010-es évek végén, 2020-as évek elején követi egy újabb. Akkor az elméleti viták hullámai a Textológiai Munkabizottságig is elértek, ennek ékes bizonyítéka a 2004-ben közreadott, már idézett *Alapelvek*.³³ Mára ezek az alapok is újragondolásra, kiegészítésre várnak. Különszámunk cikkei mind elméleti megfontolásaikkal, mind gyakorlatiasabb javaslataikkal nagyban hozzájárulhatnak ehhez, s egyúttal a számítógépes irodalomtudományról szóló diskurzust is megélénkíthetik.

³³ Szakirodalmi előzményeihez lásd: „Filológia és digitális barbárság”, ELTE BTK Bölcsészettudományi Informatika Önálló Program (BIÖP), 2004. március 4., <http://magyar-irodalom.elte.hu/biop/barbar/>.

*Számítógépes irodalomtudomány:
az adattól a transzcendentális kritikáig*

Magyarországon a digitális bölcsészet helyzete és önálló diszciplinaritása meglehetősen érdekesen alakult az elmúlt évtizedben. A magyar felsőoktatásban kérésre létezett a Debreceni Egyetemen és a Pázmány Péter Katolikus Egyetemen, hiszen néhány éves fennállása után lekerült a meghirdethető képzések listájáról. Szerencsés fejlemény, hogy 2018-tól egyre nyilvánvalóbb lett az igény az újraindításra a hallgatók és a kormányzat részéről is. Ezzel a folyamattal párhuzamosan elindult a *Digitális Bölcsészet* folyóirat az Eötvös Loránd Tudományegyetem Régi Magyar Irodalom Tanszékének védjegyével, kutatóközpontok alakultak a Szegedi Tudományegyetemen, az ELTE-n, számos konferencia szerveződik ebben a témakörben hazai és nemzetközi előadókkal. Publikációk jelennek meg továbbá a kulturális szféra digitalizációjával, valamint a digitalizáció oktatásban betöltött szerepével kapcsolatban.

Az irodalomtudományi diskurzusból is hasonló kettősség alakult ki, hiszen két véglet között helyezkednek el a tudósok: vagy a lelkesedés jellemzi a számítógépes irodalomtudományi kutatáshoz való viszonyukat, vagy egyfajta passzív rezisztencia, amennyiben nem találta relevánsnak a hagyományos irodalomtudományos kutatások esetében – a két tábor között pedig nincs érdemi diskurzus. A számítógépes irodalomtudományban leggyakrabban a stilometriai és kvantitatív alkalmazásait találhatjuk, azaz szerzőségi kérdések megválaszolását, korszakolást, kontextualizációt,¹ életműveken belüli csoportosítást,² irodalomtörténeti trendek vizsgálatát³ és ezek vizualizációját, illetve a digitális filológia kérdéseit, azaz kritikai kiadások digitális térben való megjelenítését, ennek filológiaelméleti és gyakorlati kérdéseit.⁴ Ugyan Magyarországon nem egészen jellemző, külföldön (elsősorban az angolszász szakirodalomban) számos támadás, kritika éri a digitális bölcsészetet. A támadásokban közös, hogy nominálisan a digitális bölcsészet ellen irányulnak, mégis annak csak egy ága ellen, és egyetlen irányból érkeznek. Általában ugyanis irodalomtudósok fogalmazzák meg kételyeiket, és a

¹ LABÁDI Gergely, „Az olvasó gép”, *Digitális Bölcsészet* 1 (2018): 17–34, <https://doi.org/10.31400/dh-hun.2018.1.126>.

² Az egyik legutóbbi, rendkívül izgalmas cikk ezen a téren: Kiss Margit, „Stilometriai elemzés lehetőségei magyar történeti szövegkorpuszon”, *Digitális Bölcsészet* 2 (2019): 15–33, <https://doi.org/10.31400/dh-hun.2019.2.336>.

³ Matthew L. JOCKERS, „Metaadat”, ford. LABÁDI Gergely, *Digitális Bölcsészet* 1 (2018): 83–108, <https://doi.org/10.31400/dh-hun.2018.1.242>; PÉTER Róbert, „A big data kihívás és lehetőség a bölcsészet-tudományban: digitális szövegek és metaadatok távoli olvasása”, *Magyar Tudomány* 177, 11. sz. (2016): 1323–1330, <http://www.matud.iif.hu/2016-11.pdf>.

⁴ PALKÓ Gábor, „Mit jelent a digitális filológia a szemantikus web korában? A Digiphil projektről”, *Magyar Tudomány* 177, 11. sz. (2016): 1316–1322.

digitális bölcsészetnek az irodalomtudományban betöltött szerepét kérdőjelezi meg. A kritikák a digitális bölcsészet szemére hányják „a profetikus hangvételt”,⁵ a digitális bölcsészet képviselte „neoliberális fordulatot”,⁶ egyfajta pozitivizmus visszatértét, a digitális metodológiák algoritmikus és matematikai alapjait illető problémáit. A tanulmány első felében e két utóbbi területet, azaz a pozitivizmus kérdését és a módszertant veszem szemügyre azzal a céllal, hogy egyfelől reflektáljak a támadásokra, másfelől a kritikák alapján a dolgozat második részében javaslatokat fogalmazzak meg a magyar számítógépes irodalomtudomány művelői számára.

Kezdjük az elemzést egy irodalmi idézettel, és mivel magam elsősorban Shakespeare-kutatással foglalkozom, ezért egy rövid monológgal a *Macbeth*ből.

A holnap és a holnap és a holnap:
 Nap telik nap után, mind húzza a lábát
 A jegyzett idő utolsó szaváig,
 És minden tegnap a dohos halálba
 Világított utat nekünk, bolondoknak.
 Ne éj, csonka gyertya!
 Az élet csupán sétáló árny, egy nyavalyás
 Színész, egy órát teszi-veszi magát fönn
 A színpadon és elhallgat. Egy gyagyás
 Meséje, hang és téboly,
 És nem szól semmiről. (5.5.)⁷

Ennek az idézetnek a poétikai feszültsége, a retorikai szépsége, a filozófiai mélysége az élet értelmetlenségéről mélyen megráz, valahányszor elolvasom, tanítom, színházban, filmen megnézem. A monológ megrázó mélysége számos forrásból eredeztethető, például a költői képek gazdagságából, *Macbeth*ről alkotott képünkben, a helyzetből, amelyben a szereplő elmondja ezeket a szavakat, évszázadok kritikai örökségéből. Ha azonban a számítógéppel „olvastatjuk” ezt a szöveget, valójában a szöveget szavakra szeleteljük, és ha ezeket a szavakat valamilyen szempontrendszer, például szógyakoriság szerint sorba rendezzük, a szöveg mágiája elvész. A szólistában megjelenő egyes szavak nem tudják visszaadni azt a szemantikai és hermeneutikai komplexitást, amit a beszéd jelentett. A szavak a számítógép számítási kapacitásának kihasználása végett adattá váltak.

⁵ Stanley FISH, „What’s Next for Literary Studies?”, *The New York Times*, 2011. dec. 26., hozzáférés: 2020.02.28, https://opinionator.blogs.nytimes.com/2011/12/26/the-old-order-changeth/?_r=0.

⁶ Daniel ALLINGTON, Sarah BROUILLETTE and David GOLUMBIA, „Neoliberal Tools (and Archives): A Political History of Digital Humanities”, *Los Angeles Review of Books*, hozzáférés: 2019.09.26, <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.

⁷ William SHAKESPEARE, *Macbeth*, ford. KÁLLAY Géza (Budapest: Liget Műhely Alapítvány, 2017), <https://msha.btk.ppke.hu/items/show/14>.

Ennélfogva tehát jogosnak tűnik a kérdés, hogy miként válhat adathalmazzá ez a mágikus költészet, amelyben egyszerre van jelen poétika, retorika és filozófia. Ha a szavak adattá változnak, vajon tekinthető-e irodalomnak, marad-e az költészet, ha mindennemű kontextusukat elvesztik a szavak? Hogyan beszélhetünk „grammatikai bájról”, ha az „adat” szó sterilitást, objektivitást, elvonatkoztatást hordoz magában? Szinte egyetérthetünk Marche megfogalmazásával: „Az irodalmat nem lehet értelmesen adatnak tekinteni. A probléma lényegi és nem felszínes: az irodalom nem adat. Az irodalom az adat ellentéte.”⁸ Sőt, Marche gondolatmenetét követve arra is gondolhatunk, hogy az ellentét valójában nem is irodalom és adat között húzódik meg, hanem emberi és nem emberi között. „Az algoritmusok természetüket tekintve fasisztikusak, mert az emberi viszonyoktól való különbözőség megnyugtató illúzióját keltik.”⁹ A költészet emberi, a számítógép nem emberi, sőt embertelen, és ennélfogva a számítógépnek nem lehet köze az irodalomhoz, vonhatnánk le a következtetést Marche-csal. Ezt az emberi megközelítést elvethetnénk arra hivatkozva, hogy érzelmi beállítódás, amely ellen nem érdemes érveket hozni. Mégis hasznosnak bizonyulhat ezzel a típusú hozzáállással is foglalkozni, hiszen számos esetben éppen ez a fajta érzelmi beállítódás gátolja a számítógép nyújtotta lehetőségek kihasználását az irodalomtudományban.

Érdeemes tehát leszámolni azzal a sztereotípiával, hogy az adatok a kvantitatív elemzések során a kutatást végző tevékenységétől teljesen független, mindenki számára objektív tények lennének. A szavak ugyanis abban az értelemben adatok, hogy mindegyik egyfajta absztrakciós folyamat eredménye, ti. a számítógép nem szavakat „olvas”, hanem számsorokat tapasztal a betűk helyén, végeredményként pedig 1-eseket és 0-akat, pontosabban a fázis jelenlétét és hiányát. A jelenlét és a hiány sorozatán belül kell definiálni azokat a sorozatokat, amelyeket az ember szavaknak értelmez, amely definíció számos alakot ölthet. A meghatározás először is betűsorokat illet, amelyek két üres karakter között helyezkednek el, továbbá elvonatkoztat az írásjelektől is, hiszen a számítógépnek a „holnap” és a „holnap:” különböző szavaknak tűnik, ha nem vonatkoztatunk el a központosástól. Továbbá eltekinthetünk az egyéb nem alfanumerikus karakterektől is, például a „teszi-veszi” szókapcsolat meghatározástól függően tekinthető egy vagy két szónak is. A szógyakoriság elemzésnél további absztrakciót jelent, ha nem az adott karaktersort tekintjük szónak, hanem az egyes karaktersoroknak megfelelő szótári alakot, azaz az „elhallgat” szó esetében meg kell hozni a döntést, vajon a szó a „hallgat” szóval azonosuljon, és egy igekötős igéről van szó, vagy a „csendben marad” szinonimájáról. Ugyanígy kérdés, azaz döntést igényel, hogy a szótövekhez kapcsolt ragok, jelek befolyásolják-e a számításokat vagy sem, azaz az

⁸ Stephen MARCHE, „Literature Is not Data: Against Digital Humanities”, *Los Angeles Review of Books*, hozzáférés: 2019.01.31, <https://lareviewofbooks.org/article/literature-is-not-data-against-digital-humanities/>. Az angol nyelvű idézeteket mindig a saját fordításomban közlöm, kivéve, ha külön megjelölöm a fordítót.

⁹ Uo.

„elhallgat” szó a „hall” szóként szerepeljen a listában, a „halálba” szó pedig „halál”-ként. További döntést igényel, hogy milyen szövegváltozatot használ a kutató, amennyiben ugyanannak az irodalmi műnek több szövegváltozata is létezik. Az adat tehát nem egy kutatótól és befogadótól független létező, hanem maga az elemző vagy az elemző algoritmus alkotja meg egyfajta absztrakciós folyamat segítségével. Fontos arra is reflektálni, hogy mindig egy vagy több ember áll az absztrakciós folyamatot kialakító döntések mögött, akit, akiket célok vezérnek, aki, akik előfeltevésekkel rendelkeznek. Az irodalmi szövegben a szavak mint adatok tehát semmiképpen sem steril objektivitást feltételeznek, hanem valaki által, valamilyen célból megalkotott entitásokat.¹⁰

Az ilyen értelemben létrehozott adatokat számos módon felhasználhatjuk irodalmi szövegek számítógéppel történő kvantitatív elemzéséhez. A kvantitatív elemzés arra használja fel a számítógép erejét, hogy számoljon szó- és kifejezés-gyakoriságot, más néven n-gram¹¹ gyakoriságot, n-gramok relatív gyakoriságát, szóhosszúságot, szófajok szerinti megoszlást, mondatok számát, mondathosszúságot és hasonló számmá alakítható szövegjellemzőket. Ezt a kvantitatív elemzést stilometriai, szerzőségi, filológiai kérdések felvetésére, megválaszolására lehet használni.

Bár a kvantitatív szövegelemzésnek voltak előzményei, hiszen az 1960-as évek óta használják kvantitatív elemzésre a számítógépet, talán megkockáztathatjuk, hogy a 21. században óriási felfutását tapasztalhatjuk. Meg kell azonban jegyezni, hogy a számítógép felhasználása a szövegelemzések területén már a legelső időktől az ilyen irányú kutatásokat végzőket is néha aggodalommal töltötték el, módszertani problémák vetődtek fel. Kahn 1996-ban úgy érvelt, hogy érdemes bevezetni ebbe a diskurzusba egy megkülönböztetést:

[...] megkülönböztetést egyfelől a számítógép olyan felhasználása között, hogy azt mechanikus feladatok gyorsabb elvégzésére használjuk: idézetek kikeresése, szavak megszámlálása, szólisták és szógyakorisági táblázatok készítése, esetleg statisztikai adatelemzések megjelenítése, és aközött, hogy olyan eredményeket hozunk létre, (a) amelyet eredeti tudományos eredményként értékel a tudományág, és (b) amelyet a számítógép segítségével nélkül nem lehet kivitelezni.¹²

¹⁰ Erről bővebben lásd: ALMÁSI Zsolt, „Szöveg az olvasónak: Kvantitatív módszerek és digitális Shakespeare-filológia”, *Magyar Tudomány* 177, 11. sz. (2016): 1286–1291.

¹¹ Péter Róbert így határozza meg az n-gram jelentését: „Az n-gram n darab egymást követő szóra utal: a »Magyar Királyság« például egy bigram, ha erre vagyunk kíváncsiak, akkor a kereső olyan találatokat ad vissza, ahol e két szó ebben a sorrendben áll egymás mellett.” PÉTER, „A big data kihívás és lehetőség a bölcsészettudományban...”, 1325.

¹² Abraham H. KHAN, „The Challenge of Information Technology for Literary Studies”, *Journal of Religious & Theological Information* 2, no. 2. (1996): 59–72, 59–60. Saját fordítás. A. Zs.

Ez a megkülönböztetés változatlanul fontos. A számítógép használata – még ha pontosít és gyorsabbá is teszi egyes feladatok elvégzését – nem önérték. Akkor lesz a számítógép használata ténylegesen az irodalomtudományos eszközkészlet része, ha hasznosítható tudományos eredmény megalkotásához járul hozzá, ha a tudományos közösség el is fogadja ezt az eredményt, valamint ha a számítógép használata indokolt. Ennek a meglehetősen korai kritériumrendszernek a fényében szeretném vizsgálni a számítógépes irodalomtudománynak az utóbbi években az irodalomtudomány felől érkező bírálatát.

2018 decemberében jelent meg az *Authorship* című folyóirat tematikus száma, amelynek célja a *The New Oxford Shakespeare*¹³ kritikai összkiadást kiegészítő *Authorship Companion*¹⁴ bírálatát volt. Hogy kontextusba helyezzük a Brian Vickers szerkesztette és részben maga írta tematikus folyóiratszámot, röviden bele kell tekintenünk a 21. századi, Shakespeare szerzőségét illető vitákba. Nem mintha a vita most kezdődött volna, hiszen ez visszanyúl a 19. századra, ahogy ennek természetrajzát és történetét a 20. század második feléig Shapiro alaposan feltárja *Contested Will: Who Wrote Shakespeare?*¹⁵ című könyvében. Bár a vita századok óta folyik, a 20. század vége és különösen a 21. eleje újra felveti a szerzőség kérdését, és talán leszögezhető, hogy a kérdés hozzávetőlegesen öt-tízévente újult erővel jelentkezik. Így például az *Anonymous*¹⁶ című film megjelenése, különböző képek és festmények felfedezése vagy újragondolása, esetleg stilometriai elemzések adtak erre alapot. Manapság a web 2.0-ás technológiák, a közösségi média miatt növekszik az érdeklődés, hiszen gyakorlatilag bárki fenomenológiailag azonos súllyal jelenhet meg Shakespeare-kutatási kérdésekben szakértőként, mint egy tudós. Ahogy McCrea 2004-ben szinte profetikusan fogalmaz a Shakespeare szerzőségét kétségbevonó elméletekről: „Az internet eljövételével elképzelt összeküvések soha nem látott módon terjedhetnek, és ezáltal veszélyeztetik az általunk lakott világ észlelését.”¹⁷ Ez a jelenség pontosan beleillik abba a tendenciába, hogy a világ észlelése változóban van a 21. század második évtizedében a web 2.0-ás technológiák térnyerése okán, melyre a szakirodalom gyakran mint az „igazság utáni”¹⁸ korszakra utal.

¹³ Gary TAYLOR, John JOWETT, Terri BOURUS and Gabriel EGAN, eds., *The New Oxford Shakespeare: Modern Critical Edition* (Oxford: Oxford University Press, 2016), <https://doi.org/10.1093/actrade/9780199591152.book.1>.

¹⁴ Gary TAYLOR and Gabriel EGAN, eds., *The New Oxford Shakespeare: Authorship Companion*, First edition (Oxford–New York: Oxford University Press, 2017).

¹⁵ James SHAPIRO, *Contested Will: Who Wrote Shakespeare?* (New York–London–Toronto–Sidney: Simon and Schuster, 2011).

¹⁶ Roland EMMERICH, *Anonymous*, Drama, Thriller (Columbia Pictures, 2011).

¹⁷ Scott MCCREA, *The Case for Shakespeare: the End of the Authorship Question* (Westport: Praeger, 2004), 229.

¹⁸ RÉDEY Soma, „Tudománykommunikáció az »igazságon túli« világában”, *Magyar Tudomány*, 11. sz. (2019), <https://doi.org/10.1556/2065.180.2019.11.12>.

A számítógépes stilometriai elemzések, amelyet a szerzőségi kérdések újrafelvetésére is használnak, a Shakespeare-filológiában is nagyon elterjedtek. Sorra jelennek meg közlemények, tanulmánykötetetek és monográfiák, amelyek kvantitatív, stilometriai elemzésekre épülnek, és amelyek között számos munka foglalkozik a szerzőséggel. Fontos megjegyezni, hogy a publikációk túlnyomó többsége nem azzal foglalkozik, vajon ki más lehetett a szerzője a shakespeare-i életműnek, hanem inkább azzal, hogy a kanonizált szövegek közül melyiket tulajdoníthatjuk Shakespeare-nek és milyen mértékben, illetve vajon a kétséges, apokrif iratok közül is újra lehet-e gondolni a szerzőséget. Ebbe a vonulatba illeszkednek legnagyobb hatású könyvekként Hugh C. Craig és Arthur F. Kinney *Shakespeare, Computers, and the Mystery of Authorship*,¹⁹ valamint az Új Oxford Shakespeare-kiadás fent említett kiegészítő kötete, azaz a Gary Taylor és Gabriel Egan szerkesztette *Authorship Companion* is.

KRITIKÁK

Ahogy ezek és a hasonló publikációk elhagyják a nyomdát, számos kritika is megjelenik róluk: ezen kritikák szerzői közül kiemelkedik Brian Vickers. Vickers mindig is érdeklődött a szerzőség kérdése iránt, 2002-ben például publikált egy könyvet,²⁰ amelyben megcáfolni szándékozott Gary Taylor és Donald Foster egymástól szinte független állítását (közvetve pedig Stephen Greenblatt, David Bevington Shakespeare-összkiadásait), hogy a *Funerall Elegye* (*Temetési elégia*) című költemény Shakespeare alkotása lenne. Figyelemre méltó Vickers alaposága és eltökéltsége, hiszen a felvetésre egy teljes monográfiában reagál, alaposan érvelve amellet, hogy a költemény inkább tulajdonítható John Fordnak, mint Shakespeare-nek. A részletekbe menő kritikájának egyik célpontja Donald Foster, aki a számítógépet segítségül hívva stilometriai alapon igyekezett bizonyítani Shakespeare szerzőségét. Vickers nagyon alaposan, elsősorban módszertani megfontolások alapján veszi górcső alá Foster elemzéseit, és mutatja ki hibás voltukat. A statisztikai problémák között felsorolja, hogy a Foster használta korpusz jóval kisebb volt annál, mint hogy „representatív nyelvi adatokkal”²¹ szolgálhatott volna, s hogy hiába a számolás, ha a számolt adatok tévesek. Ilyen például, hogy a versben Foster 17 hendiadys-szal számol, pedig Vickers szerint legjobb esetben is csak 2–3²² található a szövegben, valamint, hogy az adatbázisban található adatok nem publikusak, azaz nem ellenőrizhetőek. Továbbá az is befolyásolja az elemzést, hogy egyes betűsorok esetében miképpen vonatkoztatunk el a szemantikai je-

¹⁹ D. H. CRAIG and Arthur F. KINNEY, eds., *Shakespeare, Computers, and the Mystery of Authorship* (New York: Cambridge University Press, 2009).

²⁰ Brian VICKERS, *Counterfeiting Shakespeare: Evidence, Authorship, and John Ford's Funerall Elegye* (Cambridge–New York: Cambridge University Press, 2002).

²¹ Uo., 190.

²² Vö. uo., 191.

gyektől, és ezáltal mennyiben kaphatunk más és más eredményt.²³ Harmadikként pedig azt említi Vickers, hogy a mondathosszúságot illető megfontolás kétélű fegyver, hiszen az is értelmezés kérdése, mi a mondatnak tekintett egység, még inkább a korai modern angliai nyomdai gyakorlatot figyelembe véve, különösen, ha a korai modern szöveget a statisztikai elemzést végző személy modernizálja.²⁴ Hasonló probléma Foster mondathossz statisztikájával, állítja Vickers, hogy nem az átlagos mondathosszúságra koncentrálnak a szerzők, hanem a kiugró hosszúságra, pedig az előbbi az elfogadott módszer a stilometriai kutatások esetében.²⁵ Vickers továbbá sommásan megjegyzi, hogy „Foster egyáltalán nem foglalkozott az egész költemény olvasásával, hanem inkább mint nyelvi elemek halmazával, mint nyersanyaggal a számítógépes stilsztika számára”.²⁶ Ebben az észrevételben első olvasatra Marche észrevétele visszhangzik, ám itt az oppozíció nem az irodalmi emelkedettség és adatok sterilsége között tételeződik, hanem az irodalomtudományi kontextualizáció és a jelentésnélküli szétdaraboltság között.

Kevesebb, mint egy évtizeddel később Vickers két szerzőségi kérdésekkel foglalkozó könyv megjelenése kapcsán szembeállít két módszertant a szerzőségi kérdések tudományos vizsgálata esetében. Az oppozíció egyik eleme az „olvasáson alapuló” módszereket jelöli, míg a másik a „számítógép támogatta”²⁷ módszereket, amelyekről azt állapítja meg, hogy még mindig „bizonytalanok az eredményeik”.²⁸ Vickers a következő kritikai megállapításokat teszi: a művek statisztikai alapú összehasonlítását megkérdőjelezhetővé teszi az „alapvető különbség a műfajok között”,²⁹ ti. az összehasonlított művek különböző műfajokba sorolódnak, a biztos szerzőségű művek esetében tévedések találhatók, mivel ezen művek szerzősége sem teljesen megalapozott,³⁰ esetenként az eredmények összevetése a szöveg olvasásával kialakított nézettel nem szükségszerűen esik egybe.³¹ Ezen túl olyan hibákra mutat rá, mint a hagyományos kritikai irodalom figyelmen kívül hagyása,³² erősebb konklúzió levonása, mint amit a premisszák megengednek („a »hasonlóság« meghatározatlan távolságra helyezkedik el az azonosságtól”).³³ Mindezen problémák megoldását először is abban látja, hogy egy alaposan kidolgozott nyelvelméletre van szükség, amelyet a modern korpusznyelvészetben talál meg.³⁴ A korpusznyelvészeti kutatások alapján a kollokációkban, pontosabban az

²³ Uo., 194.

²⁴ Uo., 197–198.

²⁵ Uo., 199.

²⁶ Uo., 204.

²⁷ Brian VICKERS, „Shakespeare and Authorship Studies in the Twenty-first Century”, *Shakespeare Quarterly* 62, no. 1. (2011): 106–142, 111, <https://doi.org/10.1353/shq.2011.0004>.

²⁸ Uo., 114.

²⁹ Uo., 122.

³⁰ Uo., 124.

³¹ Uo., 125–126.

³² Uo., 127, 130–134.

³³ Uo., 129.

³⁴ Uo., 135.

„n-gramokban” határozza meg a megfelelő módszertant, tudniillik az egyéni, 3–4 szóból álló szókapcsolatok és a mindenki által használt szókapcsolatok különbségében.³⁵ A kollokációk keresésében két alkalmazás párhuzamos használatát javasolja: egy plágiumkereső programét, amely két szöveget hasonlít össze azonos n-gramok keresésével, valamint egy adatkereső programét, amely egy szindarabokat tartalmazó adatbázison futtatva könnyedén kikeresheti a megfelelő n-gramokat. Az ilyen módon keletkezett eredményeket, és az azokból levont következtetéseket pedig egy negatív teszttel kell megerősíteni. Cikkét azzal az optimista kijelentéssel zárja, hogy „az attribúciós kutatások új korszaka látszik elkezdődni, és a Shakespeare-kutatások sokat fognak ebből profitálni”.³⁶

Ezen előzmények után nem egészen meglepő az *Authorship* című folyóirat említett, Vickers szerkesztette tematikus száma, sem felépítését, sem pedig argumentációját illetően. Először is Vickers a bevezető cikkében különbséget tesz kvantitatív és kvalitatív kutatások között, és az utóbbi mellett teszi le a voksát, hasonlóan a 2011-es cikkéhez, ahol olvasás és számítógépes elemzés álltak szemben egymással. A kvantitatív elemzés alapvető hibájának azt tartja, hogy az

leredukálja egy komplex irodalmi műalkotás nyelvét egy ‘szavak egy táskában’-halmazra, arra a feltételezésre alapozva, hogy a gyakoriságuk megbízható szerzőségi azonosításra ad lehetőséget. A kvantitatív módszer eltekint a szövegek olvasásától, mint ahogy eltekint a jelentés megfontolásától és a szereplők egymással történő kommunikációjától is.³⁷

Ezen kritikát erősíti meg szimbolikusan a különszámot záró cikkével, ahol a kvalitatív módszertant alkalmazva cáfolja meg az Oxford csapatának téves állítását, akik egy szöveget Shakespeare-nek tulajdonítanak, holott az Vickers elemzése szerint Dekkeré.

A különszám egy másik cikkének szerzője David Auerbach, aki az oxfordiak módszertanát kritizálja részletekbe menően. Két szempontból találja Auerbach tévesnek az oxfordi módszertant, ezek a „homályosság” (*opacity*) és „a bemeneti adatok szegénysége” (*poverty of input data*). A módszertan homályosságát abban látja Auerbach, hogy szerinte nem látják el elegendő magyarázattal a módszertant, a módszertani döntéseket. A bemeneti adatok szegénysége pedig azt jelenti Auerbachnál, hogy a statisztikai elemzésekhez rengeteg adatra van szükség, míg a szerzőségi vizsgálatok esetében az egyetlen szerzőnek tulajdonítható szövegekből kinyert adatok nem elegendőek az érvényes statisztikai eredményekhez. A bemeneti adatok szegénységének másik aspektusa, hogy azok piszkosak (*dirty*), azaz sem energia, sem idő, de a kora modern szövegek esetében lehetőség sincs

³⁵ Uo., 137.

³⁶ Uo., 142.

³⁷ Brian VICKERS, „Introduction”, *Authorship* 7, no. 2 (2018): 1–10, 4, <https://doi.org/10.21825/aj.v7i2.9734>. Saját fordítás. A. Zs.

arra, hogy a szövegeket ugyanarra az absztrakciós szintre hozza az elemző. A különböző típusú szövegek, különösen az adatok nem elegendő száma miatt még bizonytalanabbá teszik az eredményeket. Auerbach egyébként két másik cikket is közölt a kötettel kapcsolatban, mindenhol részletekbe menően görcső alá véve a statisztikai modellt és az algoritmizált módszertant, és minden esetben hibásnak találta ezeket.

2019 tavaszán egy másik irányból érkező nagy port kavart kritika született meg a számítógépes irodalomtudomány (SzIT) ellen. Na Z. Da a *Critical Inquiry* című folyóirat lapjain osztotta meg kritikai észrevételeit, amelyek elsősorban a bemeneti adatokat, a módszertanokat, statisztikai modelleket illették, és amely kritikát sommásan így összegzett:

A SzIT azt állítja, hogy olyan kutatásra alkalmas eszközöket alkot meg, amelyek, még ha hibásak is, természetükből adódóan értékesek, mert a kutatás természetéből adódóan értékes. A téves besorolás önmagában lesz érdekes, a pontatlanság elmélettől alakul, a kiugró értékek esztétikai és filozófiai kutatások alapjaivá válnak, és mindez további finanszírozásra és publikációkra lesz érdemes.³⁸

A sommás összefoglalás magában hordozza az esetenként téves, ám más esetekben nagyon is helytálló kritikát. Először is, mintha azt sugallaná, a legfontosabb probléma az, hogy a számítógépes irodalomtudományt egyfajta naiv hurrá-optimizmus jellemzi. Az a fajta hozzáállás, hogy ugyan tudjuk, a felfedező eszköz, a kutatásra, elemzésre kifejlesztett alkalmazás nem jó, mégis a vele való munkálkodás, próbálkozás önértékkel bír. Ezt a tevékenységet akár lehetne felebarátian értelmezni, amennyiben a gondolatmenet úgy folytatódna, hogy az eszközök próbája a használat, és éppen a használat révén lehet a tévedéseit kiszűrni, a beállításokat módosítani, az eszközt finomítani. Da erisztikus érvelése azonban nem ebbe az irányba mutat, hanem pontosan az ellenkezőbe. Szerinte a használat nem a fejlesztést segíti elő, hanem a hibák elkendőzését azáltal, hogy a hibák valamiféle önértékké válnak: vagy egy elmélet alakul ki belőlük, vagy hatalmas esztétikai, filozófiai kutatásokra adnak alkalmat. Da kritikája tehát, még ha nem is jogos minden esetben, mégis mintha az erkölcstelen tudóst kritizálná, aki a tévedései belátása helyett inkább az előremenekülésben látja a kiutat a tévedés erényre szublimálásával. Az utolsó tagmondat tovább növeli az erkölcstelenség téziséét, hiszen a finanszírozás és a publikációk problémájára utal. Azaz, hogy ezek a kutatók az előre meneküléshez további anyagi forrásokat vesznek igénybe, és a tévedéseiket, hibás, átgondolatlan módszertanra épülő kutatásaik eredményeit tanulmányokban, cikkekben esetleg monográfiákban tárják a publikum elé. Ez pedig

³⁸ Nan Z. DA, „The Computational Case against Computational Literary Studies”, *Critical Inquiry* 45, no. 3. (2019): 601–639, 602, <https://doi.org/10.1086/702594>. Saját fordítás. A. Zs.

azt a problémát implikálja, hogy a szűkös finanszírozási keretből komoly kutatásoktól veszik el a forrásokat, a megjelentetett látszateredményeikkel megtévesztik a publikumot, illetve csak arról publikálnak, amiről valójában nem is érdemes, hiszen csak a sikertelen próbálkozást rögzítik valós eredmény nélkül.

A számítógépes irodalomtudományt illető kritikák egy generikusan eltérő, de meglehetősen fontos aspektusára szeretném végezetül felhívni a figyelmet egy elismert digitális bölcsész kutató, tehát egy belülről érkező észrevételével. Ted Underwood a hagyományos irodalomtudomány ellenérzését a számítógépes irodalomtudománnyal szemben az utóbbi retorikájában keresi. Ted Underwood így ír *A Genealogy of Distant Reading* című tanulmányában: „A retorikai bravúrokhoz szokott irodalmár közönség talán nem fogadja örömmel a társadalom-természettudományos felépítésű (módszerek, aztán eredmények, aztán konklúziók) cikket.”³⁹ Underwood kritikájában nyilvánvalóan az az igény fogalmazódik meg, hogy a számítógépes irodalomtudomány elfogadtatása az irodalomtudósokkal azon is múlik – természetesen itt abból indulunk ki, hogy a módszerek, a használt korpusz és a következtetések érvényesek –, hogy az előbbieket hogyan tárják a kutatásaik eredményeit a kollégáik elé. A kísérleti tudományok esetében ugyanis a tanulmányok úgy épülnek fel, hogy egy problémát vetnek fel a szerzők a szöveg elején, majd ismertetik mások megoldási kísérleteit, aztán egy saját hipotetikus megoldást javasolnak. Ezt követi a beszámoló a kísérletről, amelyben először is meghatározzák a mintát, az adathalmazt, az elemzés módszerét, a statisztikai módszer matematikai alapját, a komplex statisztikai módszer és egyenletek neveit. Végül a kísérlet eredményeit prezentálják táblázatokban, diagramokban, és levonják a következtetést az eredményekből. Ez a fajta felépítése a publikációknak nagy hagyományokra tekint vissza a társadalom- és természettudományokban, ám annál inkább idegen a bölcsészettudományoktól és kifejezetten az irodalomtudománytól.

Mindezen kritikákat, azaz Vickers, Auerbach, Da és Underwood kritikai megjegyzéseit figyelembe véve most arra szeretnék reflektálni, hogy miként lehet harmóniát teremteni az irodalomtudomány és a számítógépes irodalomtudomány között. Remélhetjük, hogy idővel minden magától megoldódik: ha megfelelő számú számítógépes irodalomtudományi cikk születik, egyszerűen a megszokás révén lassan elfogadottá válik mind a terminológia, mind a dolgozatok retorikai konstrukciója és eredményei. Ez talán a legegyszerűbb megoldás, és bár lehet erre számítani, talán további lehetőségek is mutatkoznak; meglátásom szerint négy aspektust különíthetünk el.

³⁹ Ted UNDERWOOD, „A Genealogy of Distant Reading”, *Digital Humanities Quarterly* 11, no. 2. (2017), <http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>, 37. bekezdés.

JAVASLATOK

Először is, talán a legegyszerűbb elkerülni azokat a szófordulatokat, terminusokat, melyeket a számítógépes irodalomtudósok az adattudománytól, információ-tudománytól kölcsönöztek, és amelyek megfelelői egyébként megtalálhatóak az irodalomtudomány szótárában, ugyanis ezek olyan asszociációkat indítanak el, mintha a számítógépes irodalomtudomány az irodalomtudománytól idegen tudományágakat követne saját megújulása, elfogadhatósága érdekében. Ha azonban mégis használni kell olyan fogalmakat, amelyek eleddig nem képezték az irodalomtudományos fogalomtár részét, akkor célszerű ezeket pontosan definiálni. Ezáltal elkerülhetővé válna továbbá, hogy olyan feltételezések sejljenek át a terminológiai választásokon, mintha objektív, az emberi tényezőtől, a hermeneutikai tértől független irodalmi adatokról vagy valamifajta determinisztikus pozitívizmusról lenne szó.

Másodszor, bár egyetértek Ted Underwood retorikai megfontolásaival, mégsem tűnik megfelelőnek a megoldási javaslata. Underwood ugyanis azt javasolja, hogy az elfogadás érdekében a számítógépes irodalomtudomány egy kicsit csaljon, amikor a kutatási eredményeiről számot ad. Szerinte „hatékonyabb lenne úgy tenni, mintha a kutatás egyszerű diszkurzív, tétel-vezérelte módon alakult volna, amit éppen egy pontdiagrammal illusztrálunk”.⁴⁰ Véleményem szerint nem az a megoldás, ha mást mutat a kutató, mint amit egyébként tenne, hanem inkább hagynia kellene, hogy az eredmények beszéljenek önmagukért, mégpedig az irodalomtudomány számára elfogadható módon. Azaz, ha a technikai részletek feltérképezése elvonja a figyelmet az eredményekről vagy azok hiányáról, biztosan nem lesz az elfogadtatás zökkenőmentes. Az is fontos továbbá, amit Vickers több ízben említett, azaz hogy a módszertan sikerrel vegyítse a hagyományos olvasáson és értelmezésen alapuló módszertant a számítógépes megközelítéssel, hiszen az irodalomtudomány számára ettől lesz hiteles mind a kutatás, mind a szerző. A számítógépes irodalomtudománynak tehát, ha elfogadásra törekszik, a két megközelítés organikus harmóniájára érdemes törekednie.

Ami az átláthatóságot illeti, három megfontolást érdemes szem előtt tartani. Először is, a korpusznak a kutatáshoz mérten elegendően fókuszálnak kell lennie, illetve a statisztika szempontjait figyelembe véve elegendően nagyoknak is, hogy a statisztikai adatok érvényesek lehessenek. Ebből a szempontból rendkívül érdekes Petr Plecháč és csapatának cikke, amelyben Shakespeare *VIII. Henrik* című darabjának szerzőségi kérdését tárgyalják. A mély tanulási szoftverrel végzett kutatás eredménye nem lett úttörő jelentőségű saját bevallásuk szerint sem, ugyanis végeredményben azt a 19. század óta tartott nézetet tudták megerősíteni, hogy a mű egy része Fletchernek tulajdonítható. Ugyanakkor ezen a paradigmán belül pontosítani tudták, melyik részt írta Shakespeare és melyiket Fletcher, illetve

⁴⁰ „It can be more effective to pretend that your work grew in a casually discursive, thesis-driven way, and then happened to be illustrated with some scatterplots you had lying around.” Uo. Saját fordítás. A. Zs.

Speddingtől eltérően a negyedik felvonás mindkét jelenetét „kevert szerzőségűnek”⁴¹ találták. A cikk maga megfelel a tudományosság számos kritériumának, pontosan adatozják milyen és honnan származó szövegekkel dolgoztak, milyen algoritmusokat alkalmaztak, hogyan szűkítették a biztosan azonosítható szövegeket a tanulási modellhez. Ugyanakkor nem reflektáltak arra, vajon négy Shakespeare-mű és négy Fletcher-darab elegendő adatot biztosít-e a tanulási korpuszhoz. Bár lehet igazuk eredményeiket illetően, mégis a korpusz mérete a kétség árnyékát veti eredményeik autentikusságára.

Az átláthatóság szempontjából fontos másodsor, hogy a korpusz hozzáférhető legyen az olvasó, a tudományos közösség számára, azaz bármikor ellenőrizhető legyen a szövegtörzs minősége, hiszen a komoly filológiai, nyelvészeti, nyelvtörténeti megfontolások nélkül megalkotott korpusz elemzéséből következő eredmények, még igazságuk esetén is tudományomódszertanilag megkérdőjelezhetők, és így valójában elfogadhatatlanok. Harmadszor pedig, ami igaz a szövegtörzshöz, igaznak kell lennie az alkalmazott szoftverre is. A számítógépes irodalomtudományos kutatások esetében nyílt forráskódú szoftvert érdemes alkalmazni, amelynek működése ellenőrizhető a felhasználás és a forráskód szintjén is. Az irodalomtudósnak ugyanis lehetőséget kell teremteni arra, hogy a saját szövegtörzs elemzése révén győződhessen meg a szoftver alkalmazásáról, valamint hogy a forráskódba betekintve, magát a kódot is szemrevételezhesse, esetleges olyan feltételezéseket keresve, amelyek az alkalmazást és ezáltal az eredményeket is befolyásolják.

Mindez hatalmas feladatot és felelősséget ró a számítógépes irodalomtudományra. A számítógépes irodalomtudósnak valójában három tudományos területen kell otthonosan mozognia: filológia, matematika-statisztika és szoftvertudományok, hiszen bármelyik területen elkövetett hiba – Brian Vickers és Da kritikáiból világosan kitűnik – a kutatás eredményeit fogja kétségbe vonni. A probléma ezzel azonban talán az, hogy kevés olyan szakember van, aki mindhárom dinamikusan fejlődő, változó tudományterületen otthonosan mozogna. Persze ebből nem szükségszerűen következik, hogy magát az ilyen jellegű kutatást fel kellene adni, hanem megoldást kell keresni a problémára. A megoldási lehetőségek közül talán járható út, ha csapatban történik a kutatás, ahol a csapatokhoz az adott szakterületek kiváló tudósai csatlakoznak, illetve hogy már a felsőoktatási tantervekbe bekerüljenek olyan tantárgyak, amelyek ebbe az irányba mutatnak. Ugyanakkor bármilyen tökéletesen összeválogatott csapat jön is létre, kritikusok mindig akadnak, hiszen minden tudományterületen belül meghozott optimális döntés, azaz a korpuszt, a matematikai-statisztikai modellt és a számítógépes architektúrát illető döntés górcső alá vehető, és egy-egy szempontból hibásnak is lehet találni. Ez a kritika azonban nem kell, hogy egy egész diszciplínát illessen, egy egész diszcipl-

⁴¹ Petr PLECHÁČ, „Relative Contributions of Shakespeare and Fletcher in *Henry VIII*: An Analysis Based on Most Frequent Words and Most Frequent Rhythmic Patterns”, *arXiv:1911.05652 [cs, stat]*, 2019. okt. 30, 9, <https://arxiv.org/abs/1911.05652>.

lína értékét meghatározza. Másfelől pedig a kritika megfontolása mindig fontos mérföldköve egy diszciplína érettségének: ha nincs tévedés, meg kell higgadtan védeni a döntést, ha pedig lehetséges egy tovább finomított módszertan alkalmazása, akkor meg kell gondolni az alkalmazását.

Az utolsó javaslatom a módszertani tudatosságot illeti, hiszen a számítógépes módszertan valójában egy új hermeneutikai szituációt teremt. Ez például abban áll, hogy amikor a drámai szöveg szavai nem együtt, hanem szólistába és táblázatokba rendeződve állnak a befogadó előtt számokkal párosítva, amelyek a szógyakoriságot jelölik, egy másik értelmezési horizonttal találkozunk. Ha például a *Macbeth* szókészletét vesszük szemügyre, egy nagyon érdekes szóra figyelhetünk fel a drámában. A leggyakrabban használt jelző a „good” (52), míg a következő jelző a gyakorisági listában a „great” (31), tehát egy másik pozitív szó.⁴² Meglepő ez azért, mert a mű világát fojtogatónak, szűknek és gonosznak találjuk, ha a drámát hagyományosan olvassuk, míg a szógyakoriságra figyelve mintha az ellenkezőjével találkoznánk, azaz a jóval. A „jó” gyakori említése a darabban utalhatna arra, hogy ez a darab nagyon is világosan az emberi dimenzió belül marad, eltérve a boszorkányok világától, ahol jó és rossz, szép és rút azonosak, értelmezhetetlen kategóriák. Az emberi világban azonban jó és rossz elkülönül, a rossz a jó ellentéte, ahol *Macbeth* pontosan tudja jó és rossz különbségét, ezért nem áztatja magát, s a rosszat rosszként választja. Ez a rossz éppen a jó ellenében, a látszólag túlcsonduló jó ellenében jön létre. A képet azonban megváltoztatja, ha a többi nagy tragédiával összevetjük ezt a számot: a *Lear király*ban a „good” 82-ször szerepel, és a következő jelző az „old” (58),⁴³ a *Hamlet*ben ismét a legtöbbet használt jelző a „good” (108-szor, míg a következő a „dead” 31-szer),⁴⁴ az *Othello*ban megint a „good” a leggyakoribb jelző (89), míg a következő az „honest” (42).⁴⁵ Ezekkel a számokkal jelentősen átalakult a kép, hiszen a többi nagy tragédiához képest a *Macbeth*ben nem is szerepel meglepően sokszor ez a szó. A számok egyelőre azonban nem egészen mondanak el mindent, hiszen a művek nem egyforma hosszúak, azaz talán többet mond a „good” szó relatív gyakorisága. Ebben a versenyt is győztes a *Hamlet* (0,00335), második az *Othello* (0,00316), harmadik a *Lear király* (0,00295) és utolsó a *Macbeth* (0,00282). Tehát a többi nagy tragédiát is figyelembe véve a „good” szó szereplése nem is olyan meglepő a *Macbeth*ben, hiszen minden más nagy tragédiában ugyanúgy a legtöbbet használt melléknév a „good”, pedig az ezekben megjelenő világ sem szép. Erre magyarázat talán az

⁴² A szöveg: William SHAKESPEARE, *Macbeth*, eds. Barbara A. MOWAT and Paul WERSTINE, hozzáférés: 2020.02.28, <https://www.folgerdigitaltexts.org/download/txt/Mac.txt>, az eszköz pedig a Voyant tools (<https://voyant-tools.org>).

⁴³ A *Lear király* szövege: William SHAKESPEARE, *King Lear*, eds. Barbara A. MOWAT and Paul WERSTINE, hozzáférés: 2020.02.28, <https://www.folgerdigitaltexts.org/download/txt/Lr.txt>.

⁴⁴ A *Hamlet* szövege: William SHAKESPEARE, *Hamlet*, eds. Barbara A. MOWAT and Paul WERSTINE, hozzáférés: 2020.02.28, <https://www.folgerdigitaltexts.org/download/txt/Ham.txt>.

⁴⁵ Az *Othello* szövege: William SHAKESPEARE, *Othello*, eds. Barbara A. MOWAT and Paul WERSTINE, hozzáférés: 2020.02.28, <https://www.folgerdigitaltexts.org/download/txt/Oth.txt>.

lehet, hogy a „good” szó általában megszólításokban szereplő bevett szófordulat, mint például a „good lord” („Jó uram”), vagy „Good night” („Jó éjszakát”) szókapcsolatokban, azaz ez egy korabeli, talán üresnek mondható szó, amely nem erkölcsi jóságról tanúskodik. Talán érdekesebb a második leggyakrabban használt melléknév ebből a szempontból, hiszen ezek mintha inkább harmóniában lennének a művekkel: *Lear király*ban az „old”, a *Hamlet*ben a „dead”, az *Othello*-ban az „honest”, míg a *Macbeth*ben a „great”. Mintha ezek a jelzők ténylegesen a tragédiák egy-egy központi gondolatára utalnának, hiszen a *Lear király* az idősödéssel is foglalkozik, *Hamlet* a halállal van elfoglalva, az *Othello* a tisztességgel, pontosabban annak hiányával, azaz a szót legtöbb esetben tévesen használják (például „honest Iago”, pedig Iagótól mi sem áll távolabb, mint a tisztesség). A *Macbeth* második leggyakrabban használt jelzője a „great”, ami leginkább egyfajta absztrakt, azaz politikai, esztétikai nagyságot jelöl, amely természetesen központi kérdése a darabnak.

Az új hermeneutikai szituáció megkövetelhet egy másik módszertant is. Általában az eredménycentrikusság miatt a módszertan ugyan fontos marad, mégis kevesebb figyelmet igényel. A módszertan, azaz hogy milyen statisztikai elméletet, milyen algoritmust használ a kutató, csak annyiban lesz érdekes egy rövid magyarázattal ellátva, hogy a befogadó megismerhesse az eredményhez vezető módszertani megfontolásokat. Ezzel párhuzamosan a módszertannal kapcsolatban fel lehet tenni a fordított irányú kérdést: ez és ez a módszertan ehhez és ehhez az eredményhez vezetett, de hogyan lehetséges, hogy a módszertan ezeket az eredményeket hozta létre? Mik azok a határfeltételek, melyek azok az előfeltevések, amelyeket a módszertan elfogad, és amelyek a módszertan alkalmazásakor pontosan ezekhez az eredményekhez vezettek? Nevezhetjük ezt a kérdésfeltevést, beállítódást Kant transzcendentális kritikája egyfajta imitációjának.

Kant kopernikuszi fordulatainak leírása segítséget nyújthat a kopernikuszi fordulatra a számítógépes irodalomtudományon belül, amely elősegítheti, megerősítheti a módszertani tudatosságot. Kant módszertana röviden és felületesen, de jelen esetben pragmatikusan, a következőkben foglalható össze. Kant a *Tiszta ész kritikájának* második kiadásának *Előszavában* a következőképpen fogalmaz: „Mostanáig feltételezték, hogy ismereteinknek mindig a tárgyakhoz kell igazodniuk; [...] Próbáljuk hát meg egyszer, nem jutunk-e messzebbre a metafizikai feladataink terén, ha feltesszük, hogy a tárgyaknak kell ismereteinkhez igazodniuk [...]”⁴⁶ A számítógépes irodalomtudomány számára persze teljesen világos, hogy egy adat és eredmény sem egy objektív igazság megmutatásában áll, azokat eleve a bemeneti adatok és a számítógépes, matematikai-statisztikai tevékenység hozza létre. Tehát sommásan úgy fogalmazhatunk, hogy a számítógépes irodalomtudomány kutatásai során valójában az adottság nyilvánvalóvá tétele, felszínre hozása helyett annak megalkotása történik meg. Ha tovább folytatjuk Kant a jelen

⁴⁶ Immanuel KANT, *A tiszta ész kritikája*, ford. Kis János (Szeged: Ictus, 1996), 34.

szempontból releváns gondolatmenetét, akkor eljuthatunk a „transzcendentális kritikához”, amelyet így határoz meg a *Bevezető*ben: „Transzcendentálisnak nevezek minden ismeretet, amely nem magukkal a tárgyakkal foglalkozik, hanem a tárgyakra irányuló megismerésünk módjával, amennyiben ez *a priori* lehetséges kell legyen”.⁴⁷ Azaz Kant amellett érvel, hogy a tárgyakat illető ismereteink valójában nem a tárgyakra önmagukban vonatkoznak, hanem arra, amit az elménk számunkra lehetővé tesz. Ha pedig ez így van, akkor a tárgyról érdemes elfordítani a tekintetet azon módra a vizsgálata irányába, amely a tárgyakat számunkra adja. Azokra az adottságokra, azaz *a priori*, tapasztalást megelőző, tiszta elvekre, amelyek egyáltalán lehetővé teszik a tárgyakkal való foglalatalkodást. Ez a tevékenység nem a transzcendentális filozófia lenne, hanem érdemesebb annak nevezni, ami: „transzcendentális kritikának, mert célja nem az, hogy gyarapítsa, hanem csupán az, hogy kiigazítsa tudásunkat, és próbakő gyanánt szolgáljon minden *a priori* ismeretünk értékes vagy értéktelen voltának eldöntéséhez”.⁴⁸ Kant szerint tehát először a transzcendentális kritikára van szükség, amely nem az ismeretek bővítéséhez vezet el, hanem csak egy eszközként szolgál annak a feltáráshoz, amit az érzékelés vagy az elme az ismeretek szerzése során már eleve hozzáad a tárgyakkhoz.

A számítógépes irodalomtudomány kopernikuszi fordulata tehát az, ha a transzcendentális kritikát a saját adottságaihoz igazítva is megvalósítja. Ez abban áll, hogy a releváns eredményeket felmutató kutatás elfordul a kutatás eredményeitől, és azt veszi górcső alá, ahogy az eredmények születnek a kutatói tevékenység során. Pontosabban azokat a lehetőségeket, azokat az adottságokat tárna fel, amelyek egyáltalán lehetővé teszik az eredményeket szinte függetlenül attól, hogy maguk a kutatás eredményei helyesek vagy helytelenek-e. Ennek a transzcendentális kritikának magába kellene foglalnia azoknak a filológiai, matematikai-statisztikai, számítástechnikai modelleknek az előfeltevés-rendszerét, amelyekből részben következik a bemeneti adattól kezdve az eredmény is. Ez azonban nem technikai jellegű elemzést jelent, hanem egy nyommal mélyebbre ásást, azt, hogy a kutatási eredményekhez vezető technológia milyen előfeltevésekre épül. Ez magában foglalja azokat az előfeltevéseket, amelyek a matematikai-statisztikai modellt, a filológia elveket, az adatbázis szerkezetét, a keresőmotor felépítését és elméletét, a jelölőnyelvet és a programnyelvet, a programot alkotják, azaz mindazt, ami összetevője annak, ami létrejön az elemzésen és az eredményeken keresztül.

⁴⁷ Uo., 69.

⁴⁸ Uo.

KONKLÚZIÓ

A javasolt elvek megoldást nyújthatnak a számítógépes irodalomtudományt illető kritikákra. A megoldási javaslatok magukba foglalják a terminológiai és retorikai finomítást, az átláthatóságot, a csapatmunkát, az új hermeneutikai szituációra történő reflexiót, valamint a kanti transzcendentális kritikának a számítógépes irodalomtudományra történő alkalmazását. Mindez elveheti a kritikák számos diszkurzív rétegét. A megfelelő reflexió a módszertani problémákat eleve megválaszolhatja. Az értetlenséget a megfelelő retorika kezelni tudja és talán egy mélyebb ellenállást is el tud simítani, amit Da úgy fogalmaz meg, hogy „mélységes összeegyeztethetlenség áll fenn a statisztikai eszközök és a tárgyak között, amelyekre alkalmazzák őket”.⁴⁹ Hasonlóan ahhoz, ahogy Vickers különbséget tett „olvasáson” alapuló és „számítógéppel támogatott” módszerek között, és ahogy Marche sommásan kijelentette, „az irodalom nem adat”. Mert Macbeth idézett beszéde nem egyszerűen egy szóhalmaz, hiszen egy szóhalmaznak nincs retorikai, poétikai és filozófiai értéke, amely ilyen röviden és pontosan meg tudja jeleníteni az élet kilátástalanságát és reménytelenségét a szavak erejével. Ez tehát nem adat. Valójában azonban ezt nem is állítja senki. Ha a fent ismertetett elvek mentén sikerül kommunikálni, talán megvalósulhat Kahn három kritériuma: a hasznosítható eredmények produkálása, a tudományos közösség elfogadása, és annak belátása, hogy az eredmény számítógép használata nélkül megvalósíthatatlan feladat. És akkor máshogy fog hangzani Macbeth beszédének első sora is: „A holnap és a holnap és a holnap.”

⁴⁹ DA, „The Computational Case against Computational Literary Studies”, 601.

GOLDEN DÁNIEL

Számítógép, szöveg, tudomány

SZÁMÍTÓGÉP

A témánkban születő írások népszerű felütéseinek egyike szerint az emberi kultúra digitális fordulatát követően annak egyetlen területe sem létezhet *számítógép nélkül* – tehát a *számítógépes irodalomtudomány* szószerkezetből a jelző mára voltaképpen el is hagyható.

A másik oldalról viszont a számítógépeknek a bölcsészettudományokban betöltött szerepe hagyományosan és még mindig kevésbé magától értetődő és kevésbé elfogadott, mint a műszaki tudományok, a természettudományok vagy a társadalomtudományok esetében. Annyi bizonyos, hogy a személyi számítógépek és a világháló használatának elterjedése óta a betű emberei is menthetetlenül belezuhantak a digitális közeg csapdájába. Persze már a kezdetektől fogva voltak olyanok, akik ebből a csapdából megpróbáltak előre menekülni: teoretikus és praktikus vizsgálatok tárgyává tenni a mélyreható átalakulást.

A nagy áttörés azonban csak a 2000-es évek közepére tehető, amikor a területtel foglalkozóknak a *digitális bölcsészet* (*digital humanities*) elnevezés alatt sikerült magukat önálló diszciplínaként elfogadtatniuk. Miközben az ezen a címen született első összefoglaló igényű kézikönyv előszavában¹ a szerkesztők is jelzik, hogy számos előzményt tudhatnak maguk mögött a *bölcsészeti informatikától* (*humanities computing*) a *számítógépes bölcsészettudományig* (*computing in the humanities*).

Amennyiben tehát a tágan értett digitális bölcsészet bő nyolcvan éves történetre tekinthet vissza, ez a folyamat jól elmesélhető a kutatói generációk, illetve az általuk követett célképzetek mentén, amelyek nagyjából együtt alakultak át azzal, hogy a számítógép műszaki vívmányból kulturális médiummá vált. A kezdeti törekvések még a hagyományos célokkal való folytonosságot hangsúlyozták:

Azt akartuk kifejezni a *bölcsészinformatika* szóval, hogy a bölcsészet továbbra is a saját kérdéseit fogja maga elé állítani, csak az őt kiszolgáló informatika legyen sokkal hajlékonyabb. Engedelmes segédeszköz legyen, a bölcsészettudománynak viszont lehetőleg ne kelljen alkalmazkodnia semmihez, így a számítógép mindenkorai lehetőségeihez sem.²

¹ Susan SCHREIBMAN, Ray SIEMENS and John UNSWORTH, „The Digital Humanities and Humanities Computing: An Introduction”, in *A Companion to Digital Humanities*, eds. Susan SCHREIBMAN, Ray SIEMENS and John UNSWORTH, xxiii–xxvii (Oxford: Blackwell, 2004).

² HORVÁTH Iván, „Digitális bölcsészet a virtuális nemzeti könyvtárban”, *Digitális Bölcsészet* 1 (2018): 121–138, 136.

A későbbiekben azután fokozatosan egyre nagyobb teret nyertek azok a megközelítések, amelyek a digitális fordulat okozta alapvető változásokkal kívántak számot vetni, párhuzamosan azzal, ahogyan az új mediális viszonyok között a helyét kereső *irodalomtudomány* saját öndefinícióját igyekezett a *médiatudomány*, illetve a *kultúratudomány* felé kitágítani.

Alighanem az első magyar nyelvű bevezetés a digitális bölcsészetbe éppen ötven évvel ezelőtt jelent meg a *Helikon* hasábjain. A téma ekkor még nem is érdemelt ki egy teljes tematikus számot, ám ha kinyitjuk ezt az 1969/3–4-es, ún. vegyes számot *A számítógépek és a humán tudományok* című összeállításnál, bizvást megállapíthatjuk: nincs új a nap alatt. Gyakorlatilag mindaz megjelenik ugyanis, amit legújabban a *digitális bölcsészet* címszó alatt szokás emlegetni. Ha beleolvasunk a Petőfi S. János szerkesztésében közölt tanulmányokba (Robert S. Wachal, Karl Kroeber stb.), valamint a *Könyvek* rovatban az ezeket magukba foglaló két gyűjteményes kötet ismertetésébe, olyan érzésünk lesz, hogy a számítógépes irodalomtudomány tekintetében a lecserélt cégér mögötti tartalom tekintetében ötven év alatt lényegében semmi sem változott: ugyanazok a mítoszok és ellenmítoszok, ugyanazok az ígéretek, relatív sikerek és relatív kudarcok köszönnek vissza.³ A számítógépek használata melletti érvek közül például a szerkesztő E. A. Bowles összefoglalójából Petőfi S. János az alábbiakat emeli ki:⁴

1. Szinte hihetetlenül rövid idő alatt képesek olyan munka elvégzésére, amelyre azelőtt egy egész életet kitöltő fáradságos munka kevés volt.
2. A számítógépet használni szándékozó humán kutatónak világosan meg kell fogalmaznia, hogy *mit* akar, mert a programozók csak így tudják elkészíteni a megfelelő programokat. Ez a tény egyrészt arra kényszeríti a kutatót, hogy következetesen végiggondolja[,] miért akarja azt, amit akar, másrészt, hogy explicit módon értelmezze fogalmait.
3. A kutatás körének mennyiségileg szinte korlátolatlan volta az általánosítás jóval nagyobb lehetőségét teremti meg, mint a kézi feldolgozás.
4. Az első eredmények elemzése alapján [...] módosítható mind a bemenő adatok megválasztása, mind a kutatási stratégia, és ez igen nagy időmegtakarítást jelent.
5. A számítógépek működését nem befolyásolják olyan tényezők, mint a kifáradás, a kutatás megszakítása vagy éppen a kutatási kedv hiánya.
6. A számítógép gyakran olyan relációkat tár fel a vizsgált adatok között, amelyekre a kutatók egyáltalán nem gondoltak.

Ismerősek a műfajok is: konkordanciák, stílusvizsgálatok, szerzőség-megállapítások, számítógépes szövegkiadás („amely különösen a kritikai kiadások előké-

³ PETŐFI S. JÁNOS, „A számítógépek és a humán tudományok”, *Helikon* 15, 3–4. sz. (1969): 372–373.

⁴ Uo., 471.

szítésénél nagy jelentőségű”), gépi szótárak. És azt is megtudhatjuk, hogy „az alkalmazások további kiszélesedését az hozhatja meg, ha megvalósul a szövegek közvetlen optikai gépi »beolvasása«.”⁵

Tagadhatatlan ugyanakkor, hogy az azóta eltelt ötven év alatt jelentős változások következtek be a számítógépes irodalomtudomány tágabb környezetét alkotó digitális kultúra területén. Ezt tükrözi a *Helikon* 2004/3-as, Kappanyos András által immár *A hipertext* címmel összeállított számában az, ahogyan a számítógépes irodalomtudományról közölt összefoglalásában Tóth Tünde határolja le választott tárgyát:

Az új médium, az internet[,] egészen új textológiai jelenségeket is hozott magával. Amikor internetes textológiáról beszélünk, akkor valójában két dologról van szó: az egyik esetben az internet pusztán a hagyományos módon készült szövegek publikálásának helye, a másik esetben viszont a textológus kutatási terepe⁶

– s deklarált módon csak az előbbiről ír. Az itt közölt fordítások (Aarseth, Bolter, McGann, Simanowski) már a digitális közeg újfajta kihívásaival szembesítenek, a valódi szemléleti paradigmaváltást ugyanakkor a *Helikon* 2014/3-as *Archívum*-száma jelzi, ahol ez utóbbira irányuló munkálatokra a textológia kiterjesztett új neveként megjelenik a *médiaarcheológia* – amely nagyrészt megismétli azokat a felismeréseket, amelyeket már az 1980–90-es évek elméletírói is felfedeztek.

A különbség a megítélésben van: a korai könnyed optimizmust a körültekintő pesszimizmus váltotta fel. Utóbbi a két világ ütközéséből fakad: az állandóságot adó megőrzésre épülő, ezért statikus hagyományos bölcsészettudományok és a folyamatos forgalmazásra törekvő digitális kultúra radikálisan eltérő értékrendje és mechanizmusai között. A számítógép, amely valaha az emberileg elérhetőnél nagyságrendekkel nagyobb rögzíthetőséget ígért, mára sok tekintetben a rögzíthetlenség totalizálását hozta el.

De a technológiai fejlődés ipari valósága sem váltotta be a hozzá fűzött humanista reményeket. A Text Encoding Initiative kezdeményezői már 1990-ben azt panaszolták, hogy a bölcsészettudományok digitális kiterjedését az alábbi körülmények akadályozzák:⁷

- nem oszthatunk meg dokumentumokat a kollégáinkkal, mert eltérő szövegszerkesztő programokat használnak;
- a kiadók nem szedhetnek egyenesen a lemezeinkről, hanem mindent újra kell gépelniük új hibákat bevezetve;

⁵ Uo., 472.

⁶ Tóth Tünde, „Online kritikai szövegkiadás Magyarországon az ezredfordulón”, *Helikon* 50, 3. sz. (2004): 417–440, 420.

⁷ Steven J. DeRose, David G. Durand, Elli Mylonas and Allen H. Renear, „What is Text, Really?“, *Journal of Computing in Higher Education* 1, no. 2. (1990): 3–26.

- a campus túlsó végén lévő nyomtató más (vagy egyáltalán semmilyen) eredményt produkál, mint az előcsarnokban lévő;
- az online munkával töltött évek során felhalmozott hatalmas mennyiségű szövegben (tanulmányok, feljegyzések, könyvlisták stb.) még mindig nem közlekedhetünk hatékonyan: tudjuk, hogy egy idézet ott van *valahol* a lemezünkön, és mégsem vagyunk képesek megtalálni.

Az anomáliákra általuk adott válasz a tartalmi és formai szabványosítás növelése – miközben a folyamatok jó része továbbra is az ellenkező irányba halad.

Az a dinamika, amivel a digitális média mint minden korábbi formát önmagába felszívni képes metamedium a kulturális tartalmak teljesen szabad, korlátok nélküli újrakeverését teszi lehetővé, alapvetően ellentmond az írás kultúrtechnikájához a legutóbbi időkhöz tapadó stabilitásnak. Ez a tapasztalat jelenik meg a *Helikon* 2017/2-es *Hálózatelmélet és irodalomtudomány* számában Kirschenbaumtól és Druckertől közölt írásokban is. A jelenség maga pedig a klasszikus filológia és bölcsészet számára egyaránt fenyegető vízióként jelentkezik, ahogy arra már egészen korán Nyíri Kristóf figyelmeztetett: „A hagyományos bölcsészettudományok tárgya s ugyanakkor előfeltétele az írott, a szilárd alakot nyert szöveg.”⁸

SZÖVEG

Ami elől semmiképpen nem térhet ki a jövő bölcsészettudománya, az tárgyának minőségi változása, vagyis az anyag természetének átalakulása. A digitális kultúra korának alapmetaforája a szöveg; a digitális bölcsészet az univerzális szöveg elmélete, avagy a szöveg univerzális elmélete. Minden szöveg, még a kép is, amennyiben olyan diszkrét szekvenciákból áll, amelyek manipulációjára irányuló eljárások a szövegtudományok hagyományos módszereivel jól leírhatóak.⁹

Az elektronikus szöveg korai elméletírói (Nelson, Kittler, Bolter, Heim, Landow, Lanham, Dyson, Aarseth) igyekeztek a technológiai potenciált a teoretikus potenciállal párosítani. A következő generáció által már „utópikus felbolydulásnak” minősített elképzelés hívei az új kommunikációs közegektől a „nyomtatott formátumokban latens lehetőségek realizálódását”¹⁰ remélték a szövegek sokszorozhatóságának, összekapcsolhatóságának, átjárhatóságának és audiovizuális kiterjesztésének lehetőségeit illetően. Előképeiket az orosz formalistáktól a francia strukturalizmuson át az amerikai dekonstrukcióig találták meg, olvasataik pedig a foucault-i értelemben *antihumánusak* voltak, és persze nagyon is *humánusak*, amennyiben a szövegértelmezés évezredes paradigmáján belül ma-

⁸ NYÍRI Kristóf, „Bölcsészettudományok az írásbeliség után”, *Világosság* 37, 6. sz. (1996): 3–16, 3.

⁹ HORVÁTH Iván, *Gépeskönyv* (Budapest: Balassi Kiadó, 2006), 144–146.

¹⁰ Vö. JOHANNA DRUCKER, „Miről értesít az (im)materialitás: a szöveg mint kód az elektronikus környezetben”, ford. MEZEI GÁBOR, *Helikon* 63, 2. sz. (2017): 281–301.

radtak – értsd: egy adott szöveg korát nem fizikai vagy kémiai analízis révén kívánták megállapítani, hanem az *olvasás* különféle módjait működtetve.

Az elmúlt három évtized ismételten bebizonyította, hogy a kultúra konzervatív rendszerként működik. A nyilvánosság ollója szétnyílt: egyik oldalról az intézményesült formák, amilyenek a bölcsészettudományok is, nehezen mozdulnak, ragaszkodnak jól bevált mechanizmusaikhoz, a másik oldalról a tömegek hátra sem nézve hagyják maguk mögött az alfabetikus alapú kultúrát. A két világ közötti távolság és feszültség egyre nő, s a digitális bennszülöttek felnövésével valódi földindulásszerű átalakulások következhetnek.

Mindeközben a digitális kor információs túltermelésének mennyiségi nyomása alól a humaniorák kétféle minőségi irányba próbálnak menekülni: az egyik az *adatbázis*, a másik a *hipermédia* műfaja. A szövegek kulturális terének radikális átalakulására adott bölcsészettudományi válaszok ilyen értelemben alapvetően két csoportba sorolhatóak attól függően, hogy az új jelenségek értelmezési keretét a géppel támogatott olvasás *statisztikai* vagy *asszociatív* lehetőségeit jelölik ki.

A szöveg mint adatbázis

A bölcsészettudományok lényegét mindig is a szövegek értelmezése adta; ez az, ami mindenkori képviselőit a végtelenített információáramlás közepette továbbra is érdekelné fogja. Tudományos fogalmi rendszerezésre ezen belül a jeltömegben megfigyelhető *mintázatok*¹¹ lesznek érdekesebbek – Franco Moretti szerint például az időben hosszan változatlan minták adhatják az alapvető szövegformákat, a rövidebb időtávon megfigyelhetőek pedig a műfajokat.¹² A feladat minden esetben a szabályosan visszatérő sajátosságok azonosítása lesz, amiben a gépi intelligencia valóban felbecsülhetetlen szolgálatot tehet a humán intelligencia számára, amennyiben teherbírása és monotonitáé, azaz mennyiségi és minőségi megbízhatósága összehasonlíthatatlanul jobb azénál. Ugyanakkor az is vitathatatlan, hogy a mintázatok *elkülönítése* és *megnevezése* mindenképpen kreatív emberi hozzájárulás eredménye. Az ún. jelölőnyelvek sem tesznek mást, mint hogy a szövegfolyam bizonyos részeit önálló jelentéssel bíró egységként azonosítják, s az adott szegmenshez egy szintaktikai kódolási rendszer révén egy adott jelentést társítanak.

Ez éppen az *adatbázis* legegyszerűbb értelmezése. A hozzáadott jelentések és kategóriáik önálló entitásként kezdenek létezni, azaz az alapszöveg olyan információkkal gazdagodik, amelyek önmagukban is értelmes rendszert alkotnak, s más szövegekre is vonatkoztathatóak. A kritikai kiadás, a konkordancia, a kvantitatív stilisztikai elemzés, a szerzőségi kérdések vizsgálata, a műfajok elhatárol-

¹¹ Vö. SYI [SZAKADÁT István], „Olvasni az oszlopok között”, in *Csomópontok: A digitális kultúra jellemzői és egymásra hatásuk*, szerk. RAB Árpád, 128–153 (Budapest: Gondolat Kiadó–INFONIA–Óbudai Egyetem Digitális Kultúra és Humán Technológia Tudásközpont, 2017).

¹² Franco MORETTI, *Distant Reading* (London: Verso, 2013), 86, 24. lábjegyzet.

lása nem más, mint a természetes nyelvi szöveg elemeinek önmagukon túlmutató jelentésekkel való felruházása, azaz *szisztematikus interpretáció*. Így értendő tehát az a kitétel, amely szerint „[a] számítógépet itt többnyire arra használják fel, hogy a szubjektív megfigyeléseket megerősítsék vagy megcáfolják.”¹³

A naiv megközelítés valamiféle új objektivitást is remél ettől, de ezt a digitális bölcsészek többsége határozottan elutasítja. Ahogyan Labádi Gergely figyelmeztet: jóllehet a gépi támogatás igénybe vétele

korábban elképzelhetetlen mértékű kontextualizációt tesz lehetővé, amely az egyedi szövegek környezetének, az egyes szövegeket meghatározó irodalmi, társadalmi, kulturális trendeknek az azonosítását, elemzését ígéri [...] Ugyanakkor nagy mennyiségű adat számítógépes elemzése egyszerűen csak egy alternatív módszer adatokat gyűjteni, hipotéziseket felállítani, ám ezeket ugyanúgy kritikával kell kezelni, és értelmezni szükséges.¹⁴

S ezzel cseng össze végkövetkeztetése is:

amíg a magyar írott örökség kapcsán nem kezdődik meg egy szisztematikus, egységes átírási, modernizálási elvek alapján történő, nyílt hozzáférésű digitalizálás, addig az ellenőrizhetőség, megismételhetőség alapvető tudományos kritériumainak sok vizsgálat nem fog tudni megfelelni, a *distant reading* módszerei közül sokat nem lehet biztonsággal használni.¹⁵

Nem véletlen, hogy a bölcsészettudományokkal kapcsolatba kerülő informatikai szakemberek ezt általában alapvető problémaként, az ilyen jellegű alkalmazások fejlesztésének komoly gátjaként említik: „Az adatgyűjtéssel kapcsolatos legkülönfélébb nehézségek elsősorban az irodalomtudományi kontextusban releváns adatok nagyfokú bizonytalanságából, az adathiány kiküszöbölhetetlenségéből és az adatok elkerülhetetlen interpretálásából adódnak.”¹⁶ „Ezek a problémák nagyban hátráltatják és sok esetben meg is akadályozzák a tudásalapú szövegkezelés alkalmazását, az így megvalósítható informatikai szolgáltatások kialakítását, az ezekre épülő kutatások elvégzését, így végső soron új tudományos eredmények elérését.”¹⁷

Ám ha az adatkapcsolatok nem használhatóak szigorú értelemben, akkor metaforikusak lesznek – amint erről maga Moretti is beszámol a *Hamlet* szereplőinek

¹³ PETŐFI S., „A számítógépek és a humán tudományok”, 472.

¹⁴ LABÁDI Gergely, „Az olvasó gép: Berzsenyi Dániel versei távolról”, *Digitális Bölcsészet* 1 (2018): 17–34, 19.

¹⁵ Uo., 34.

¹⁶ BENKE Gábor, „Informatikai alkalmazások az irodalomtudományi kutatásban: Módszertani és gyakorlati problémák”, *Literatura* 29, 1. sz. (2003): 98–106, 99.

¹⁷ MÉSZÁROS Tamás, „Mit nyújthat a modern informatika az irodalomtudomány számára?”, *Magyar Tudomány* 177, 11. sz. (2016): 1310–1315, 1314.

kapcsolati hálóját vizsgáló kutatását ismertette. „A súly és az irány vizualizálásának azonban semmilyen módját nem találtam, amely ne volna esetlen; ennél fogva az e tanulmányban szereplő hálózatok mind kézzel készültek”, mivel „ezek kis hálózatok, ahol az intuíció még szerepet játszhat; olyanok, mint a hálózatelmélet gyermekkorra az irodalom számára; rövid boldogság a statisztika rideg felnőtt-kora előtt”.¹⁸ Majd következőképpen vonja le az eset tanulságát:

Visszanézve az elkészült munkára, nem nevezném hibának ezt az irányváltást: végül is a hálózatelmélet *valóban* segített újradefiniálni a cselekmény elméletének pár kulcsszempontját, amely az irodalomtudománynak fontos aspektusa. [...] Nem, nem volt szükségem a hálózatelméletre; de a hálózatokra valószínűleg igen. [...] Amit a hálózatelméletből merítettem, inkább *vizualizáció* volt, semmint a fogalmak: annak lehetősége, hogy kinyerjük a karaktereket és az interakciókat a drámai struktúrából, majd jelhalmazzá alakítsuk azokat, ami egy pillantásra áttekinthető egy kétdimenziós térben. Alapvetően ugyanúgy használtam (jól vagy rosszul) az elméletet, ahogy a kartográfiát az *Európai regény atlaszában* vagy a diagramokat a *Gráfok, térképek, fákban*: mint az irodalmi adatok elrendezésének egy módját, ami előfeltételezett egyfajta rendező elvet – de nem egy teljes fogalmi építményt.¹⁹

Hasonló értelemben figyelmeztet Stephen Ramsay is a naiv szcientizmus veszélyére:

Az ember nem érti, miért van szükség arra, hogy ezekről a belátásokról úgy beszéljünk, mintha az interpretációtól független, objektív módon jönnének létre, s így módon alkalmasak volnának a hagyományos bölcsészeti jelentésképzés cáfolatára. Mintha minden retorikai tárgy volna – az 'adatokat' kivéve.²⁰

Az általa javasolt algoritmikus irodalomtudomány (*algorithmic criticism*) módszertani szigorúságra törekszik az alkalmazott eszközök vonatkozásában, ugyanakkor a hangsúlyt nem a valóság leképezésének igényére, hanem a módszertani döntések átláthatóvá tételére helyezi. Kevésbé tartja fontosnak a vizsgálatok megismételhetőségét, sokkal inkább az átalakíthatóságukat, továbbfejleszthetőségüket. Az így dolgozó irodalomtudósnak ugyanúgy tudatában kell lennie saját olvasói eljárásai előfeltevésszerűségének, mint például egy feminista olvasat létrehozójának – mindössze a digitális bölcsész ennek során a számítógép segítségével is támaszkodik. Az algoritmikus kritika valahol középen helyezi el magát a pozíti-

¹⁸ Franco MORETTI, „Hálózatelmélet, cselekményelemzés”, ford. KASZAI Máté, *Helikon* 63, 2. sz. (2017): 216–257, 217.

¹⁹ Uo., 227.

²⁰ Stephen RAMSAY, *Reading Machines: Toward an Algorithmic Criticism* (Urbana: University of Illinois Press, 2011), 5.

vizmus álmának beteljesítését remélő és a számítógép közreműködését teljesen elutasító irodalomtudós között. Az emberi kultúra termékeire általánosságban úgy tekint, mint amelyek radikálisan *átalakíthatók, újrendezhetők, szétszerelhetők és újra-összeszerelhetők*.

Ez indokolja James Dobson kezdeményezését a digitális bölcsészet kritikai fordulatát²¹ illetően, amelynek kapcsán ő is azt hangsúlyozza, milyen távol áll valójában a számítógépes irodalomtudomány (*computational criticism*) a megcélozni vélt empirikus-objektív tudományoszménytől. Véleménye szerint a gépi támogatottságú olvasás variációi a kulturális analitikától²² a kvantitatív formalizmuson, illetve távoli olvasáson,²³ az algoritmikus irodalomtudományon²⁴ és makroanalízisen²⁵ át a számítógéppel támogatott szövegbányászati kizárólag az által nyerhetik el bölcsészeti jelentésségüket, ha világossá teszik a gépi olvasás során alkalmazott eljárásaikat – éppen úgy, ahogy ezt manapság a hagyományos olvasatok esetében elvárjuk.

A szöveg előkészítése (az információk rendezése) a gépi olvasásra ugyanis minden esetben emberi előfeltevések mentén történik, vagyis a mintázatok illesztése, az egyedinek az általánoshoz rendelése mindig esetleges, azaz *vitatható*. Ahogy a digitalizálás sem más, mint az analóg folytonos fenomenális világ lefordítása diszkrét jelek rendszerére – az ennek során alkalmazott szegmentáció, normalizáció, klasszifikáció, klaszterezés és modellezés valamennyi mozzanata értelmezésre és kritikára szorul, mivel ezek mindegyike voltaképpen dekontextualizációs és rekontextualizációs műveletek sorozata.²⁶

Dobson ennél is továbbmegy, amikor átveszi Barbara Herrnstein Smith kritikáját, miszerint az, amit például Matthew Jockers csinál az általa választott szövegbázis kapcsán, már nem is irodalomtudomány, sokkal inkább *társadalomtudomány*. A Moretti által megénekelte „irodalom vágóhídján” áldozatul esett szövegtömeg emancipálása deklaráltan nem az immanens esztétikai érték rekonstruálása jegyében történik, sokkal inkább valamilyen tágabban értett kultúrszociológia érvényesítésével. S ez alighanem fontos határvonalra hívja fel a figyelmet: ma már nem az választja el a bölcsészettudományokat a társadalomtudományoktól, hogy vizsgálataikhoz igénybe veszik-e a gépi intelligencia támogatását vagy sem, hanem az, hogy *milyen kérdéseket* tesznek fel a humán és/vagy gépi intelligencia szá-

²¹ James E. DOBSON, *Critical Digital Humanities: The Search for a Methodology* (Urbana: University of Illinois Press, 2019).

²² Lev MANOVICH, „Cultural Analytics: Visualizing Cultural Patterns in the Era of More Media”, *Domus*, no. 923. (2009): 112–115.

²³ MORETTI, *Distant Reading*.

²⁴ RAMSAY, *Reading Machines...*

²⁵ Matthew L. JOCKERS, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013).

²⁶ Ennek megvalósítását lásd a vélhetően első magyar nyelvű adatbázis-kritikában: TóTH Tünde, „Irodalomtörténészek a Bábeli Könyvtárból: A régi magyar vers repertórium”, *Irodalomtörténeti Közlemények* 101, 1–2. sz. (1997): 133–145.

mára. Vagyis az ugyanazon korpuszon ugyanazon technológiai eljárásokkal végzett számítógépes szociológiai, pszichológiai, nyelvészeti, irodalomtudományi vizsgálatokat kizárólag a *kérdésirányok* eltérő volta alapján különböztethetjük meg egymástól.

Mindebből következően Dobson a digitális eszközökkel támogatott olvasás Katherine Haylestől származó népszerű felosztásával (*close – machine – hyper*)²⁷ szemben a Ramsay egyik esszéjében²⁸ szereplő két alaptípust, a keresést (*searching*) és a böngészést (*browsing*) fogadja el. A kettő közti különbséget pedig abban látja megragadhatónak, hogy míg az előbbi „irányított, és a keresést végző bizalmat táplál az intézményi és szervezeti struktúrák iránt”, addig az utóbbi „anarchikusan játékos, unalomtól motivált”, voltaképpen tudatos „kitörés a katalógusok, mutatók és besorolási állományok általi korlátozás alól az új felfedezések érdekében”.²⁹

A szöveg mint hipermédia

Ha a strukturalista megközelítésben a kép is szöveg, akkor a hermeneutikában a szöveg is kép.³⁰ A szöveggént azonosított képpontok összessége az emberi észlelés számára először egészként adott, s csak tanult szabályrendszerek révén történik meg a kép szegmentálása szövegegységekké, majd ezekhez az egységekhez az önálló jelentések hozzárendelése.

A digitális közegben a képpontok ráadásul interaktívvá is tehetőek, ami azt jelenti, hogy kapcsolóként definiálva érintésre egy másik mediális tartalmat emelnek be az észlelésbe. Ezzel a hagyományos *statikus szöveg* a tapasztalásban *dinamikus hipermédiává* válik, aminek eredményeként számos lehetőséggel bővül: keresések és találati listák; felhasználó által alakítható (testre szabható) megjelenítés, komplex audiovizuális élmények stb. (Persze mindezek mögött szintén a szöveg „adatbázissá” bővülése húzódik meg, amennyiben adott pontjai valós időben önállóan hivatkozhatóak és manipulálhatóak lesznek a további tartalmakkal való összekapcsolódás érdekében.)

A digitális kor alapvető jellegzetessége a tartalmak közti ugrálás közben felépülő jelentésháló, amely nincs tekintettel a művek, műfajok, médiaformák közti határookra. A szöveg többé már nem lehet tisztán alfabetikus jelhalmoz, hanem különböző jeltípusok egyvelege lesz. Az „olvasás” és „írás” elsajátítandó eljárásai nem korlátozódhatnak a betűsorok felfejtésére, hanem „kritikailag informált írásbeliségként kiterjesztendőek a grafikus dizájnrá, vizuális narratívára, időalapú

²⁷ N. Katherine HAYLES, *How We Think: Digital Media and Contemporary Technogenesis* (Chicago: Chicago University Press, 2012), 73–74.

²⁸ Stephen RAMSAY, „The Hermeneutics of Screwing Around; or What You Do with a Million Books”, in *Pastplay: Teaching and Learning History with Technology*, ed. Kevin KEE, 111–120 (Ann Arbor: University of Michigan Press, 2014).

²⁹ DOBSON, *Critical Digital Humanities...*, 35.

³⁰ HÁRTÓ Gábor, „A grafikai mozzanat a szövegben”, *Literatura* 21, 2. sz. (1995): 204–212.

médiára és az interfészek kialakítására”.³¹ A képernyőkorszak térben és időben kiterjesztett szövegeinek a technikájában és az esztétikájában is óhatatlanul megjelenik a film- és videókultúra. A képkivágás, a megvilágítás, a kép- és hangvágás készségei észrevétlenül beépülnek a digitális közegben történő szövegalkotába.

A digitális közegben azonban még a teljesen hagyományos szöveg helyzete is átalakul. Annak idején a különböző technológiai környezetben különböző észlelési tapasztalatként jelentkező elektronikus szöveg önmagától való elkülönülődésének tapasztalata érthető módon váltotta ki a változatlan mélystruktúra és a változó felszíni megjelenés közti különbségtétel igényét; a tudományos feladatnak az előbbi megragadása, azaz a *látszat* mögött megbúvó *lényeg* azonosítása és bemutatása tűnt. A TEI megalkotóit ez a filozófiai ábránd vezérelte: a szöveg nem más, mint tartalmi elemek vagy jelentések hierarchikus rendje.³²

Csakhogy a szövegstruktúrák analizálása és rögzítése fölött a napi gyakorlatban jelen állás szerint győzelmet arattak a szöveget *képként* kezelő alkalmazások. Mégpedig alapvetően azért, mert ez a mód adja a legkevesebb áttétellel, azaz a legegyszerűbben és leggyorsabban az emberi észlelés számára ismerős élményt. A nyomtatott könyv korában az alfabetikus átírás egyértelműen könnyebben forgalmazható volt, mint a hasonmás. A technológiai fejlődés mára ezt megfordította – a szöveges információk továbbítására szolgáló alkalmazások, illetve formátumok elsődlegesen képiek, s csak másodlagos funkcióként igyekeznek támogatni a betűalapú kereshetőséget. A felhasználói igények mentén ilyen praktikusán vágta át a szolgáltatásipar az *egymást átfedő hierarchiák*³³ teoretikus gordiuszi csomóját.

Másfelől arra a kérdésre, hogy mit jelent ezeknek a jelentéseknek az explicitte tétele, azaz felszínre hozása és objektív rögzítése, egy radikális szövegelmélet fényében az a válasz adható, hogy se többet, se kevesebbet, mint az adott szöveghez kapcsolt újabb szövegeket, ahol a kapcsolatok magyarázatának feladatát megint csak újabb szövegek látják el. A digitális közeg ma az elemek kapcsolódásának ez a végtelen burjánzása uralja, s nem a precízen felépített klasszikus adatbázisok, amelyek komoly hátrányokkal indulnak versenybe a digitális kor olvasójának figyelméért: nem látnak a mélyükre a keresők, nem férhetőek hozzá egyszerűen, s mivel nem könnyen migrálhatóak, hosszabb távú fenntartásuk a változó technológiai környezetben már-már reménytelen.

A digitális bölcsészet hatókörébe tartozó elektronikus szövegek szokásosan megkülönböztetett két alapvető típusa a *digitalizált* (azaz nyomtatott előzményből azzal a szándékkal létrehozott, hogy az olvasási tapasztalatban azonosnak mutakozzon) és a *digitális* (azaz eleve elektronikus közegben, eredeti képződményként

³¹ Anne BURDICK, Johanna DRUCKER, Peter LUNENFELD, Todd PRESNER and Jeffrey SCHNAPP, *Digital Humanities* (Cambridge, MA: The MIT Press, 2012), 10.

³² Lásd: DE ROSE, DURAND, MYLONAS and RENEAR, „What is Text, Really?”.

³³ Allen RENEAR, Elli MYLONAS and David DURAND, „Refining our Notion of What Text Really Is: The Problem of Overlapping Hierarchies”, in *Research in Humanities Computing*, eds. Nancy IDE and Susan M. HOCKEY, 263–280 (Oxford: Oxford University Press, 1996).

létrejött). Amire azonban igazán szükség lenne, az ezekhez képest egy harmadik lehetőség, amikor a nyomtatott előzményből a digitális változatot azzal a szándékkal hozzuk létre, hogy az olvasási tapasztalatban a nyomtatottól eltérő, a digitális közegből fakadó hozzáadott értékkel bírjon (minimum a megjelenítés eloldása a nyomtatott formátumtól, maximum a kereshetőség stb.). A legnagyobb kihívást pedig a mátrix negyedik tagja tartogatja, azaz a digitálisan született források tudományos digitális kiadása. Amennyiben ragaszkodni kívánunk a filológusi munka hagyományos lépéséhez, ez utóbbi megvalósítása sem történhet másként, mint a működésben lévő szöveg szabad szemmel történő megfigyelésére építve a jelentésstruktúra értelmezésével, majd kódolásával. Ennek a hozzáadott értéknek a megjeleníthetősége azonban a digitális közeg előnyeit kihasználó szövegek esetében még egyáltalán nem látszik. Mert hogyan is kellene kinéznie az utókor kulturális emlékezetére áthagyományozni kívánt tudományos elektronikus kiadásnak példának okáért a magyar digitális kultúra hőskorát meghatározó olyan szövegtermékek esetében, mint az *ABCD*³⁴, a *Gólem*³⁵, a *Disztichon Alfa*³⁶ vagy a *Répaköltészet*?³⁷

A nagy kérdés tehát az, hogy a könyv „fogságából” kiszabadulni látszó szöveg a hipermédia közegébe átörökítődve képes lesz-e megőrizni hagyományos értékeit, a *kritizálhatóságot* és az *archiválhatóságot*. Avagy feloldódik a folytonosan jelenlévő virtuális audiovizualitás archiválhatatlan és kritizálhatatlan állandó jelenidejűségében, amely esetben – hasonlóan a helyhez és időhöz kötött művészetekhez, mint amilyen a tánc vagy a színház – a történeti vizsgálat számára csak a *jelenségek emlékezete* lesz hozzáférhető, a jelenség maga nem.

TUDOMÁNY

Ami a bölcsészettudományok alapjait érintő nagy átalakulással kapcsolatos igazán feszítő kérdéseket³⁸ illeti, azokra az eltelt évtizedek alatt sem születtek megnyugtató válaszok. Másként fogalmazva: nem történt érdemi előrelépés a kultúra digitális fordulatának bölcsészettudományos kezelését illetően. Vannevar Bush hetvenöt évvel ezelőtti vízióját újra elolvasva bosszantóan egyértelmű, még

³⁴ *ABCD interaktív magazin*, 1994–1996, hozzáférés: 2020.02.28, <http://syi.hu/abcd.html>.

³⁵ FARKAS Péter, *Gólem*, hozzáférés: 2020.02.28, <http://www.interment.de/golem/index.htm>.

³⁶ PAPP Tibor, *Disztichon Alfa*, hozzáférés: 2020.02.28, <https://mek.oszk.hu/11700/11744/pdf/>.

³⁷ *Répaköltészet*, <https://forum.index.hu/Article/showArticle?t=9015877>, valamint *Prae* 3–4. sz. (2000), hozzáférés: 2020.02.28, http://magyar-irodalom.elte.hu/prae/pr/200002/222_repakolteszet.html.

³⁸ Lásd például: Seamus Ross, „Az emlékezet jövője: Digitális megőrzés és társadalomtudományok”, [ford. n.], *Világosság* 42, 7–9. sz. (2001): 31–57; Daniel L. GOLDEN, „The Electronic Turn: Changes in Textual Structure”, in *Vers une nouvelle philologie*, ed. Levente SELÁF (Hallgatói Információs Központ, 2007), <https://www.tankonyvtar.hu/hu/tartalom/tkt/vers-une-nouvelle/ch13.html>.

mindig mekkora a távolság a tudományos kutatás lehetséges jövője³⁹ és valóságos jelene⁴⁰ között.

Ami egyértelműnek látszik, ahogy azt a digitális pusztulás példái is megerősítik, hogy McLuhan szellemében nem elég a *tartalommal* foglalkozni, muszáj a *formával* is, különben baj lesz a *hozzáférréssel*. „Ez azt jelenti, hogy a »mű« többé már nem pusztán a tartalom, hanem *minden*: a környezet, amelyet a mű előadására és publikálására alakítottak ki; az interfész és az adatszerkezet, a háttér-adatbázis és a kód, amely a befogadói bevonódás többféle formáját teszi lehetővé.”⁴¹

Az arra irányuló kérdés pedig, hogy ezzel a kihívással szembesülve miként alakulnak át a bölcsészettudományok több száz- vagy ezeréves hagyományokkal bíró gesztusai a kulturális tárgyak rögzítésére (módosítására, kiegészítésére, szerkesztésére), tárolására, feldolgozására és hozzáférhetővé tételére vonatkozóan, úgy is feltehető: *hogyan juthatunk el az irodalomtudomány szolgálólányaként értett digitális filológiától a digitális bölcsészet valódi emancipációjáig?*

Digitális filológia

Csábító lehet az a megközelítés, amelynek értelmében a digitális bölcsészet a hagyományoshoz képest pusztán léptékváltást képvisel: hiszen konkordanciák, írói szótárak léteztek jóval a számítógépek feltalálása előtt is, ami tehát az elmúlt évtizedekben zajlott, az voltaképpen nem más, mint a jól ismert folyamatok sebességének és mértékének ugrásszerű növekedése.⁴² Legalább ekkora súllyal van jelen az az álláspont, amely szerint a digitális eszközök révén a filológiában a tőlük függetlenül is létező teoretikus kihívásokra már korábban adott válaszok (lásd *genetikus textológia, új filológia*) kaptak felbecsülhetetlen értékű támogatást a technológia oldaláról, hogy elméleti vízióikat a gyakorlat világába is áttemelhesék.⁴³

A helyzet azonban bizonyos értelemben sokkal rosszabb. Ugyanis a hagyományos bölcsészet szemléletéhez a végsőkig ragaszkodó módszertannal készített szövegkiadások mint artefaktumok sem menekülhetnek a mediális átalakulás csapdájából, mivel a digitalizálódó szövegkezelési eljárások elkerülhetetlen alkalmazása során óhatatlanul digitális dokumentumok formáját öltik.⁴⁴ Ebből követ-

³⁹ Vannevar BUSH, „As We May Think”, *The Atlantic Monthly* 176 (1945): 101–108.

⁴⁰ HAYLES, *How We Think...*

⁴¹ BURDICK, DRUCKER, LUNENFELD, PRESNER and SCHNAPP, *Digital Humanities*, 127.

⁴² Roberto A. BUSA, „Foreword: Perspectives on the Digital Humanities”, in SCHREIBMAN, SIEMENS and UNSWORTH, *A Companion to Digital Humanities*, xvi–xxi.

⁴³ Lásd: PALKÓ Gábor, „Digitális filológia: számítógép anyaszerepben”, *Filológiai Közlöny* 61, 2. sz. (2015): 187–199.

⁴⁴ A digitális közegben termelődő szövegek archiválásának nehézségeiről általában lásd: Matthew G. KIRSCHENBAUM and Doug RESIDE, „Tracking the Changes: Textual Scholarship and the Challenge of the Born Digital”, in *The Cambridge Companion to Textual Scholarship*, eds. Neil FRAISTAT and Julia FLANDERS, 257–273 (Cambridge: Cambridge University Press, 2013); DRÓTOS László és KOKAS Károly, „Webarchiválás és a történeti kutatások”, *Digitális Bölcsészet* 1 (2018): 35–53; Thorsten RIES and Gábor PALKÓ, „Born-digital archives”, *International Journal of Digital Humanities* 1, no. 1. (2019): 1–11.

kezően a legortodoxabb célkitűzések mentén szerveződő vállalkozásoknak is korábban soha nem látott munkaszervezési kihívásokkal kell megküzdeniük, amennyiben a digitális közegben (is) meg kívánnak jelenni:

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 kiadások ma már csoportmunkában készülnek, s a teljesítmény nem feltétlenül elkülönült részekben ölt testet, amelyek így névhez köthetőek lennének. Egy kiadás adatbázisában a szöveg hagyományos textológiai-filológiai feldolgozása és kódolása más és más személyhez kötődhet, 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.⁴⁵

De a legalapvetőbb kérdésekre sem látszik egyértelműen, hogy mi lenne a helyes válasz. Például az egyik TEI szabványból a másikba részben gépi konverzió útján, részben humán szakértői közreműködéssel történő átkódolást, amelyhez megjelenítés viszont nem áll rendelkezésre – „tekinthetjük-e ezt egyáltalán kiadásnak?”⁴⁶ Ami természetesen arra a dilemmára vezethető vissza, hogy „a kiadást lezárt egészként képzeljük-e el (mint ami ki van adva)”⁴⁷.

Az elméleti kihívásoknál látszólag kevésbé súlyosak a szövegek forgalmazásának gyakorlatában jelentkező nehézségek; valójában ezek ássák meg végleg a hagyományos bölcsészet sírját. Az adatbázis-szemléletű digitális bölcsészet eredménye ugyanis a „túl okos szöveg”, ami a kiadói és archívumi gyakorlatban inkább teher, mint előny: „A kiadó előtt álló új kihívás a szövegtest és a jegyzetek létrehozása egy adott szövegkiadás számára a kutatási szakasz adatbázisaiból és jelölőnyelvi struktúráiból álló komplex rendszer korrekt konverziós lépések sorozatával végrehajtott szisztematikus redukciója révén.”⁴⁸

⁴⁵ DEBRECZENI Attila, „Az elektronikus kiadás színe és visszája”, in *A humán tudományok és a gépi intelligencia*, szerk. TOLCSVAI NAGY GÁBOR, A humán tudományok alapkérdései 3, 48–61 (Budapest: Gondolat Kiadó, 2018), 58–59.

⁴⁶ MARÓTHY Szilvia, „A tudományos elektronikus szövegkiadások feldolgozása”, in *Valóságos könyvtár – könyvtári valóság: Könyvtár- és információtudományi tanulmányok 2018*, szerk. KISZL Péter és Csík Tibor, 351–357 (Budapest: ELTE Könyvtár- és Információtudományi Intézet, 2018), 354.

⁴⁷ MARÓTHY Szilvia, „Elektronikus szövegkiadások a könyvtárban”, *Tudományos és Műszaki Tájékoztatás* 64, 6. sz. (2017): 298–309, 299.

⁴⁸ GÁBOR KECSKEMÉTI, „Electronic Textual Criticism: A Challenge to the Editor and to the Publisher”, in *New Publication Cultures in the Humanities: Exploring the Paradigm Shift*, ed. Péter DÁVIDHÁZI, 91–98 (Amsterdam: Amsterdam University Press, 2014), 94.

Tulajdonképpen ennek tudható be az is, hogy miközben a magyar digitális filológia gyakorlata világszínvonalúnak mondható,⁴⁹ a hazai könyvtárak mint kulturális archívumok alapvetően felkészületlennek mutatkoznak a digitális filológia által létrehozott kulturális termékek befogadására. „Mind a PIM OPAC, mind a DEENK WebPac felől közelítve az említett tételek teljességgel hiányoznak, tehát az egyes szövegkiadásoknak, melyek egy összetett szolgáltatást alkotnak, nincs katalogikus feldolgozása annak ellenére, hogy a metaadatok elméletileg könnyen kinyerhetők lennének a TEI Headerből.”⁵⁰

Amennyiben egy digitális filológiai projekt komolyan veszi mindazokat a kihívásokat, amelyekkel a digitális média sajátosságai szembesítik,⁵¹ s kísérletet tesz egy olyan információs ökoszisztéma felépítésére, amely naprakész és adekvát válaszokkal szolgál az adatminőség, az adatkonzisztencia, az adatbiztonság, a hosszútávú megőrzés, a hozzáférhetőség terén, akkor ennek a feltételrendszernek a megteremtése olyan komoly infrastrukturális és ebből következően finanszírozási terhekkel jár, amelyek csak stabil intézményes háttérrel teljesíthetők. Ez utóbbi ugyanakkor a valóságban a legritkább esetben bizonyul adottnak, az itthoni és a nemzetközi szinten még mindig inkább elszigetelt magánkezdeményezésekkel találkozunk, semmint pontosan kidolgozott állami stratégiák mentén szabályozott, sokoldalú intézményi együttműködésekkel. Nem függetlenül attól a tudomány-szociológiai paradoxontól, amely szerint a megvalósíthatóság és a fenntarthatóság egyaránt együttműködést kívánna, miközben a humán tudományok művelői definíció szerint még mindig elsősorban egyéni babérokra törnek. A szellem magányos hőségének képét nem könnyű feladni az összetett alkalmazkodást kívánó digitális vállalkozások kedvéért, a gyakorlatban ezért gyakran maradnak az egy személyes zárványok, amelyek lényegében készítőikhez tapadnak, s sorsuk az alkotó személyes sorsának alakulásával együtt változik.

Az új, digitális szöveg tehát új, digitális filológiát követel meg, amint arra elsősorban Matthew Kirschenbaum⁵² hívja fel a figyelmet. Ahogy a hagyományos filológia figyelme szükségképpen kiterjed a szöveg hordozóira és azok sajátosságaira a papiruszoktól a kódexeken át a nyomtatott könyvig, úgy a hiteles digitális bölcsészettnek készen kell állnia arra, hogy megbirkózzon a digitális hordozók és közegek jelentette speciális kihívásokkal. Ez a digitális archeológia ki kell terjedjen az eszközök és tárolók működési elveinek vizsgálatára, adott esetben rekonst-

⁴⁹ A teljesség igényével készült áttekintéseket lásd: TÓTH, „Online kritikai szövegkiadás Magyarországon az ezredfordulón”; HORVÁTH, „Digitális bölcsészet a virtuális nemzeti könyvtárban”; MARÓTHY Szilvia, „Tudományos szövegkiadások a hálózaton: Áttekintés”, *Irodalomtörténeti Közlemények* 122, 5. sz. (2018): 617–633.

⁵⁰ MARÓTHY, „Elektronikus szövegkiadások a könyvtárban”, 301.

⁵¹ Lásd: PALKÓ GÁBOR, „Mit jelent a digitális filológia?”, *Magyar Tudomány* 177, 11. sz. (2016): 1316–1322, 1319.

⁵² MATTHEW G. KIRSCHENBAUM, *Mechanisms: New Media and the Forensic Imagination* (Cambridge, MA: The MIT Press, 2007); MATTHEW G. KIRSCHENBAUM, „Ancient Evenings: Retrocomputing in the Digital Humanities”, in SCHREIBMAN, SIEMENS and UNSWORTH, *A Companion to Digital Humanities*, 185–198.

ruálására vagy restaurálására is. A szöveg- és kultúratudományok mellé tehát felzárkózik a „szoftver tanulmányok” (*software studies*), amely a digitális kor kulturális interfészeinek kritikai vizsgálatát tűzi ki céljául, mint amelyek alapvetően meghatározzák a befogadói élményt.⁵³

A mediális összetettség következménye a hozzáférés erőteljes technológiai kiszolgáltatottsága, ami radikális különbség a könyvkultúrához, de még az elektronikus kommunikáció előző korszakához képest is. A különböző nyomdák által előállított könyvek, a különböző rádióállomások és tévécsatornák által sugárzott műsorok ugyanazon befogadói eljárások egyszeri elsajátításával hozzáférhetővé váltak. A digitális média kínálta megjelenítésbeli komplexitás ára, hogy az alkalmazott technológiák, illetve az ezek születését és működését kísérő társadalmi és gazdasági folyamatok olyan fokú komplexitást mutatnak, amely jelen állás szerint uralhatatlannak bizonyul. Leegyszerűsítve: az elmélet szintjén triviális egységesítési és szabványosítási igények rendre vereséget szenvednek a piaci verseny logikájával szemben – az eszközök fejlesztésétől kezdve a digitalizálási projektekig bezárólag.

Ezen a ponton a digitális korszak filológiájának egy súlyos paradoxonnal kell megküzdenie, amennyiben ellentmondásba kerül egymással a *megőrzés* és a *feldolgozás* évezredes parancsa. Az egyszerű szövegrögzítő rendszerek szegényes lehetőségeket kínálnak a jelentések rögzítésére. Az összetett rendszerek gazdagokat, ám a használhatóság kérdése korántsem egyértelmű: minél összetettebb a kódolási rendszer, annál összetettebb dekódolásra van szükség. Minél komolyabb, elmélyültebb, azaz tudományosabb igényű a feldolgozás, annál összetettebb eljárások, eszközök használatát feltételezi, ami a létrejövő digitális dokumentumot az előbbiek tekintetében *szélsőségesen egyedivé* teszi. Ami egyszersmind azt is jelenti, hogy minél strukturáltabb, annál több feltételnek kell teljesülnie a hozzáféréshez is, amit a befogadónak a saját környezetükben ki kell alakítaniuk. Mindez együtt pedig az önmagát villámsebességgel felülíró digitális környezetben egyértelműen a gyors elmúlás záloga.

A digitális kultúra nagy adatvesztései abból származnak, hogy a dekódoló rendszerek pótolhatatlannak (azaz elvi és gyakorlati szinten is rekonstruálhatatlannak) bizonyulnak. A nyomtatott könyv mint évszázadokon át lényegében változatlan kódolási rendszer alkalmas volt a bölcsészettudományok felépülésére és megszilárdulására. A digitális kódolási rendszerek jelenlegi elavulása ennek az ellenkezőjével jár, s ebben a tekintetben az összetettebb rendszerek (és az azokban kódolt szövegek) egyelőre lényegesen rosszabbul járnak, mint az egyszerű kódolásúak.

A paradoxon tehát így szól: minél komplexebb, annál sérülékenyebb. Erről tanúskodik a magyarországi digitális filológia története, ezen belül is legemblematikusabb darabjainak sorsa. Ezek a kiadások saját élettörténettel rendelkeznek – Ma-

⁵³ Lásd: Lev MANOVICH, *The Language of New Media* (Cambridge, MA: The MIT Press, 2001); David M. BERRY and Anders FAGERJORD, *Digital Humanities: Knowledge and Critique in a Digital Age* (Cambridge–Malden, MA: Polity Press, 2017).

gyarországon minden bizonnyal a leghosszabbal a *Répertoire de la poésie hongroise ancienne (RPHA) / A régi magyar vers repertórium*,⁵⁴ amely már öt-hat migráción esett át váltakozó eredménnyel, s hosszú távú fenntartása a mai napig nem biztosított.⁵⁵

Digitális bölcsészet

Úgy tűnik tehát, hogy a bölcsészettudományoknak annyiban kellene átalakulniuk, amennyiben tárgyuk, a szöveg átalakul. Ez idáig választhatunk a régi szövegek új módszerrel történő vizsgálata és az új szövegek régi módszerrel történő vizsgálata között – miközben az új szövegek új módszerekkel való vizsgálata még mindig várat magára.

Másfelől viszont ha valóban az a helyzet, hogy a kultúra digitális fordulatát követően lényegében nem létezik többé a digitális eljárásoktól tisztán tartott bölcsészettudomány, amennyiben valamennyi hagyományos munkafázis: a forrásokhoz való hozzáférés, a szövegek létrehozása, a szerkesztés és a publikálás egytől egyig digitális eszközök kisebb-nagyobb mértékű igénybevételével történik – akkor mi indokolja az önállóságot követelő *digitális bölcsészet* elkülönítését?

Az öt szerző együttműködéséből, bevallottan programadó céllal, *Digital Humanities* címen született kötet⁵⁶ által adott válasz szerint az, hogy az erre felesküdt szakemberek a tudományos kutatás engesztelhetetlen híveiként nem elégszenek meg az átlagos felhasználónak nyújtott lehetőségekkel, hanem újabb és újabb kihívások elé állítják önmagukat, a feldolgozandó anyagot és a rendelkezésre álló eszközöket. Ameddig ez az avantgárd viszonyulás a technológiához fennáll, addig van értelme a digitális bölcsészetet megkülönböztetni az általános bölcsészettől – ami viszont definíció szerint csak addig érvényes, amíg a kísérleti eredmények,⁵⁷ vagyis az adott fejlesztés nem válik általánosan elterjedté.

A digitális bölcsészet története e szerint megírható annak történeteként is, hogy mikor milyen technológiai kísérletek voltak napirenden, amelyek aztán beépültek a hétköznapi használatba. Ennek jegyében a tudományos írásbeliség új formáit kérhetjük rajta számon, például az *interaktív tanulmányokat*, az *elektronikus gyűjteményeket* vagy a *digitális narratívákat*.⁵⁸ Az adatbázisok útvesztőiben való bolyongás helyett a szó legmélyebb értelmében felhasználóbarát megoldásokra van szükség: „Ne tévesszük szem elől a fő célt. Azért írunk, *ma* azért írunk irodalom-

⁵⁴ Lásd: HORVÁTH, „Digitális bölcsészet a virtuális nemzeti könyvtárban”, 135.

⁵⁵ Jelenleg a címlap itt: <http://rpha.elte.hu/>, a használati útmutató itt: <http://magyar-irodalom.elte.hu/repertorium/dokument/index.html>, az adatbázis pedig itt: <https://rpha.oszk.hu/> érhető el (hozzáférés: 2020.02.28.).

⁵⁶ BURDICK, DRUCKER, LUNENFELD, PRESNER and SCHNAPP, *Digital Humanities*.

⁵⁷ Vö. GOLDEN Dániel, „Tudományos írásbeliség és infokommunikációs technológia”, *Világosság* 47, 2. sz. (2006): 11–16.

⁵⁸ William G. THOMAS III, „The Promise of the Digital Humanities and the Contested Nature of Digital Scholarship”, in *A New Companion to Digital Humanities*, eds. Susan SCHREIBMAN, Ray SIEMENS and John UNSWORTH, 524–537 (Chichester: Wiley-Blackwell, 2016), 533.

történetet, hogy narratívákat adjunk a digitális örökséggel nyakon öntött felhasználónak”.⁵⁹ Az olvasás a digitális örökség teljességéből értelmes egységeket kihasználó bejárású utak kérdése lesz – s ebben a helyzetben ezek létrehozása lehet a digitális kor irodalomtudósának feladata.⁶⁰

Ahogy William G. Thomas III mondja: elérkezett az ideje az igazi bölcsészettudománynak. A digitalizálás, gyűjteményképzés és kiadás gesztusait követnie kell a sokrétű elemző munkának. Ezeknek a munkálatoknak a „digitálisan születő értelmezői tudományhoz” (*digitally native interpretive scholarship*) kell vezetnie. Ha van a digitális bölcsészettudománynak önmagával szembeni adóssága, akkor ez az: hiába születtek rendkívül izgalmas „hibrid” tudományos munkák, amelyek egyesítik magukban az archiválást, a sajátos eszközök kialakítását, a kommentárok készítését, az adatok gyűjtését és vizualizációját,⁶¹ ezek elenyésző számban kaptak megfelelő kritikai figyelmet. Mivel pedig ez utóbbi a bölcsészettudomány elidegeníthetetlen lényegéhez tartozik, mindez azt a tévképzetet keltheti, hogy a digitális bölcsészettudomány nem hozott létre igazán számottevő eredményeket.

A *Digital Humanities* szerzői szerint a fenti értelemben vett kísérletező olvasatok programja melletti elköteleződés óvhatja meg attól is a digitális bölcsészettudományt, hogy gyakran önmaga számára is váratlanul a társadalomtudományi kutatások területére lépjen. A statisztikai vizsgálatokat muszáj hermeneutikai elemzéseknek kísérniük – csak így biztosítható a bölcsészettudományok legnemesebb hagyományainak továbbélése.

Hiszen ez a humaniorák lényege: a *humán értelem* által a kollektív kulturális emlékezet elemei közti kapcsolatok létrehozása az értelmezői közösségek fenntartásának céljából. És ez jelöli ki a digitális bölcsészettudomány határait is: ha majd egyszer a fenti képletben a humán értelmet gépi értelemre kell cserélnünk – nos, ennek a maga teljességében *gépi olvasásnak* a mibenlétéről már csak az a bizonyos *gépi értelem* szolgálhat majd adekvát beszámolókkal...

Amíg azonban még belátható ideig humán projektek uralják a terepet, addig a bölcsészettudományok művelésének egyre kevésbé perifériális, egyre inkább centrális alkotóelemévé válik a digitális filológia, ami egyben visszatérés a filológia klasszikus szerepéhez: a textológia segédtudományból első tudománnyá válik.⁶² A digitális kultúra idáig részletezett gyengeségeire (sérülékenység, intézményesíthetlenség stb.) a *Digital Humanities* manifesztuma a következő választ adja: *mindig új kísérletek kellenek!* Az így létrejövő diszciplínát egyenesen generatív bölcsészettudománynak (*Generative Humanities*) nevezik, ahol a jelző a módszertan, a számítási kapacitás és a tudásreprezentáció új módjainak kombinációjára utal, amelynek jellemzői a termékeny kudarcok prototípusokon, teszteléseken keresztül történő

⁵⁹ HORVÁTH, „Digitális bölcsészettudomány a virtuális nemzeti könyvtárban”, 133.

⁶⁰ HORVÁTH Iván, „Számítógép és irodalomtudomány”, *Alföld* 61, 3. sz. (2010): 9–28.

⁶¹ THOMAS, „The Promise of the Digital Humanities...”, 525.

⁶² Jerome MCGANN, „Coda: Why digital textual scholarship matters; or, philology in a new key”, in FRAISTAT and FLANDERS, *The Cambridge Companion to Textual Scholarship*, 274–288, 275.

keresése, s az a megközelítés, amely szerint minden megoldás új kérdéseket szül, és ez éppen a fejlődés záloga. Hiszen a kudarc elfogadása az innovációs ökoszisztéma kiküszöbölhetetlen része. S ezek a folyamatok nemcsak szövegeket generálnak (elemzéseket, kommentárokat, narratívákat, kritikákat), hanem képeket, interakciókat, többféle médián átívelő korpuszokat, szoftvereket és platformokat.

A valódi kísérletezés nélkülözhetetlen velejárója a kudarc: a digitális bölcsészet minden látszat ellenére nem az örökkévalóságnak, hanem a pillanatnak dolgozik. Ennek tökéletes példáját olvashatjuk az *ELTE BTK Bölcsészettudományi Informatika Önálló Program* működésének Horváth Iván által megvont mérlegéhez fűzött lábjegyzetben:

Az is előfordult, hogy semmi újat nem találtunk ki, de túl korán vettünk át külföldi ötleteket. A GNU Public Licence-ét használtuk, nem sejtve, hogy a későbbi Creative Commons lesz a nyerő. Bevezettük az *f-bookot* ($f = \text{free}$), nem sejtve, hogy az Európai Unió az Open Access mellé fog állni. Ennek összes haszna csak annyi lett, hogy hálózati kiadványaink még nem kaptak ISBN-t (ilyesmire valaha kérni sem lehetett), és nem szerepelnek MTMT-bibliográfiáinkban.⁶³

A digitális bölcsészet „egy olyan univerzumot kutat, amelyben a nyomtatott szöveg nem a tudás létrehozásának és terjesztésének elsődleges közege többé”.⁶⁴ Nem azzal foglalkozik, hogy „a nyomtatott szöveg modelljeit egyszerűen digitális öntőformákba töltse, hanem a tudás *print-plus* és *post-print* modelljeinek kifejlesztésével.”⁶⁵ Ez többet jelent a tudásközvetítő rendszerek frissítésénél; a humán területek kognitív és episztemológiai újjáalakítását is magában foglalja: „A *print-plus* korszak képernyői és kiterjesztett valóságai lehetővé teszik a korpuszok rétegezését, szűrését és verziózását az egyetlen repozitóriumon belüli sokféle bejárési út együttes létezése, az érvelés multilineáris formái érdekében.”⁶⁶ Ebből következően a digitális bölcsészet hangsúlyai is áthelyeződnek: a bölcsészeti informatika kezdeti hullámainak hagyományos szövegek központúságától (a szógyakorisági vizsgálatoktól a szövegelemzéseken és jelölőnyelvi kódolásokon át a hiperszövegek szerkesztéséig és szöveges adatbázisok létrehozásáig) elmozdulva a tudás létrehozásának és szervezésének grafikus módszerei, a vizualitás, a dizájn szerepének felismerése felé.

Nyilvánvaló, hogy a fenti követelményeknek megfelelni akaró bölcsészettudomány művelése olyan sokrétű szakértelmet kíván, amelynek birtoklása nemigen várható el az egyes embertől. Együttműködésre van tehát szükség, a magányos

⁶³ HORVÁTH, „Digitális bölcsészet a virtuális nemzeti könyvtárban”, 135, 53. lábjegyzet.

⁶⁴ BURDICK, DRUCKER, LUNENFELD, PRESNER and SCHNAPP, *Digital Humanities*, 122.

⁶⁵ Uo., 125.

⁶⁶ Uo.

kutatás (*solitary study*) helyét a sokféle szakértelmet egy adott projekt érdekében egyesítő *kutatócsoport* veszi át.

A *projekt* egy olyan tudományos kutatás (*scholarship*), amelyhez szükség van dizájnrá, menedzsmentre, tárgyalásra és együttműködésre.⁶⁷ Ez a forma persze csak a bölcsészettudományokban tud az újdonság erejével hatni, a természettudományokban, műszaki tudományokban már rég nem. (Természetesen léteztek munkacsoportok korábban is, például a nagy kézikönyvek összeállítása esetében, a nagy újdonság itt az interdiszciplináris összetételben van, amely a szakzsargonok közti fordítást teszi szükségessé.) A szerzők éppenséggel azt állítják, hogy a digitális bölcsészeti projektek a szerzőségi modellek tekintetében a természettudományos laboratóriumi kutatásokkal, illetve a csoportos előadó-művészeti produkciókkal lesznek analógak.⁶⁸ Mindez egy sokrétűen strukturált, hierarchizált rendszert adhat ki a vezető kutatótól a *crowdsourcing* és *citizen science* jelenségig.⁶⁹ A projekt áttöri a bölcsészettudomány ezeréves életidegenségét és elszigeteltségét, amennyiben együttműködéseket kell kezdeményeznie az adott kutatási témában érintett szakmai szervezetekkel, testületekkel, különböző memóriaintézményekkel (könyvtárakkal, múzeumokkal, levéltárakkal).

Mindez szükségképpen magával hozza a bölcsészképzés átalakításának igényét is:

A digitális bölcsészeti projektekbe bevont hallgatók, nagyon hasonlóan a természettudományos laboratóriumhoz, a gyakorlat révén tanulnak ebben a kiterjesztett osztályteremben szakértők vezetése alatt, a társaikkal együttműködésben. Míg a hagyományos bölcsészettudomány a készségek megszerzését az eredeti eredményeket hozó kutatómunka előzetes feltételének tekintette, addig a digitális bölcsészet felgyorsítja ezt a gyakornokoskodást azzal, hogy a hallgatókat a kezdetektől a kutatói közösség részévé teszi.⁷⁰

A pedagógiai modell érdekesen találkozik a kutatásról alkotott vízióval, amely szintén a kreatív, generatív és kísérletező folyamatokat részesíti előnyben a produktumokkal szemben, a verziókat és a kiterjeszthezőségeket a tökéletes kiadásokkal és a kutatási jelentésekkel szemben.⁷¹

Amennyiben teret engedünk annak a felismerésnek, hogy a kultúra digitális fordulata nyomán az emberiség előtt álló egyik legfontosabb kihívás a kulturális örökség átvándoroltatása a nyomtatottból az elektronikus közegbe, a digitális bölcsészet a változatlan örökségként értett kultúra eme válságára adott válasznak

⁶⁷ Uo., 124.

⁶⁸ Uo., 125.

⁶⁹ Michael NIELSEN, *Reinventing Discovery: The New Era of Networked Science* (Princeton: Princeton University Press, 2012).

⁷⁰ BURDICK, DRUCKER, LUNENFELD, PRESNER and SCHNAPP, *Digital_Humanities*, 126.

⁷¹ Uo., 22.

is tekinthető. A bölcsészettudományok képviselői hagyományos szereptudatuk birtokában mindig is amellett törnek majd lándzsát, hogy ha ezt a feladatot az archívumi szakemberek vagy a kulturális piaci szereplők végzik el, a minőség mérhetően gyengébb lesz, mint ha a humanista tudósok mai leszármazottaira bíznák azt. Merthogy ez az *átírás* a szó valamennyi értelmében *kreatív* gesztus, nagyjából úgy, ahogyan a középkori kódexmásolók esetében. S amit a fentiek fényében végképp nem szabad szem elől tévesztenünk, hogy ez a bizonyos átírás nem egyszeri aktus, hanem a papíralapú adathordozók korához képest összehasonlíthatatlanul nagyobb gyakorisággal, voltaképpen *folytonosan* jelenlévő feladat. Ennek a soha véget nem érő migrálásnak az elvégzése elkötelezett kolostorokat és szerzeteseket kíván. Mindez különleges küldetéstudattal töltheti el a digitális kor új humanizmusát,⁷² egyszersmind új jelentéssel ruházhatja fel a *philologia perennis* fogalmát.⁷³

⁷² Jerome McGANN, *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction* (Cambridge, MA: Harvard University Press, 2014).

⁷³ McGANN, „Coda...”.

INTERJÚ

hic Rhodus, hic salta: *Tito Orlandi és Julianne Nyhan beszélgetése*

Az alábbi interjú Olaszországban, Rómában készült, 2014. október 17-én, körülbelül 9 órakor. Orlandi felidézi, hogy a legkorábbi, számítógéppel kapcsolatos emléke az 1950-es évekből származik, ekkor látott először IBM készüléket egy milánói IBM-üzlet kirakatában. 1960 körül PhD-konzulensével, Ignazio Cazzanigával együtt kísérleti jelleggel igyekeztek felmérni, hogyan lehetne Ágoston *Isten városa* című művének kritikai kiadását a lyukkártyás technológiával segíteni. Az 1970-es években kezdett alaposabban számítástechnikával foglalkozni, amikor a kopt kéziratokról általa felhalmozott rengeteg információ kezelésének gyakorlati problémájába ütközött. Már a kezdetektől tisztában volt a számítógépes megközelítések lehetséges korlátaival: korai találkozása Silvio Ceccato munkásságával óvatosságra intette a kibernetikai megközelítésekkel kapcsolatban. Az alkalmazott matematikus Luigi Cerofolini munkái voltak rá a legnagyobb hatással a módszertani kérdések megértésében; egyebek mellett ő tanította meg neki a Unix rendszert. Az elmélettel kapcsolatban kihangsúlyozza, mekkora hatással volt rá Turing univerzális számítógépének megértése. A modellezés számítógépes bölcsészeten belüli fontosságát hangsúlyozó munkája¹ valójában megelőzte McCartyét.² A digitális bölcsészet eredetével kapcsolatos öröklött hiedelmek, különösen a Fr. Roberto Busa S. J. szerepét övező mítoszok megkérdőjelezésén kívül az interjúban Orlandi amellet érvel, hogy a digitális bölcsészet nem figyel eléggé a számításméleti alapokra.

ÉLETRAJZ

Tito Orlandi 1940. június 18-án született Cremenóban (Como). 1963-ban diplomázott a Milánói Egyetemen (Università degli Studi di Milano), disszertációját az ókori filozófia történetéből írta. 1976-tól 2010-ig az olaszországi „La Sapienza” Università degli Studi Roma egyetem kopt nyelv és irodalom professzora volt. 1992-től 2010-ig az egyetem Centro Interdipartimentale di Servizi per l’Automazione nelle Discipline Umanistiche (CISADU, Bölcsészettudományi Automatizá-

¹ Lásd erről pl. Tito ORLANDI, „The Scholarly Environment of Humanities Computing: A Reaction to Willard McCarty’s talk on The computational transformation of the humanities”, Home Page: Prof. Tito Orlandi, [s. d.], hozzáférés: 2020.03.02, <http://www.cmcl.it/~orlandi/mccarty1.html>.

² Willard MCCARTY, *Humanities computing* (Basingstoke–New York: Palgrave Macmillan, 2005).

lasi Szolgáltatások Központja) egységének igazgatója volt. Emellett (1984 és 1994 között) ő igazgatta az egyetemén tartott korai, a számítógépes bölcsészet területével foglalkozó *Informatica per le Scienze Umanistiche* kurzust is. Jelenleg is igazgatója a Corpus dei Manoscritti Copti Letterari (CMCL) projektnek, amely egy sokféle forrásból álló, elsősorban a kopt nyelv és az egyiptomi keresztény kultúra i. e. 2. és i. e. 12. század közötti történetére koncentrálnó online tudományos gyűjtemény. A kopt tudományok területén elért számtalan érdemén kívül kiemelkedő színvonalú és úttörő jelentőségű munkát végzett a számítógépes bölcsészet olaszországi és azon túli elterjesztéséért.³ Társszerzője a *Computing in Humanities Education: A European Perspective*⁴ című kiadványnak. Tiszteletére Fiormonte és Perilli 2011-ben Festschriftet állított össze.⁵

INTERJÚ

Julianne Nyhan: Mi a legkorábbi emléke a számítógépekkel vagy a számítástechnikával kapcsolatban?

Tito Orlandi: Az 1950-es években láttam az IBM készülékeit a milánói üzletük kirakatában; ekkor jöttem rá, hogy valami ilyesmi létezik. Még gimnazista voltam.

Utána, még első egyetemi éveim alatt megismerkedtem Silvio Ceccatóval. Mond ez a név Önnek valamit? Valószínűleg nem, de ő volt az egyik első olasz értelmiségi – egyben persze a milánói egyetem professzora is –, aki érdeklődött a mesterséges intelligenciák, illetve a mesterséges érvelés előállításának módszertana iránt.⁶

Ekkoriban filológiát tanultam, és a filológia professzorommal együtt megpróbáltunk kialakítani egy olyan rendszert (ez 1960 körül történt), amellyel azt akartuk felmérni, elkészíthető-e Szent Ágoston *Isten városa* című művének latin kritikai kiadása ilyen lyukkártyás számítógépek segítségével. Bizonyos értelemben én voltam a projekt kezdeményezője, mivel én beszéltem róla a filológia professzoromnak, Ignazio Cazzanigának. Kíváncsi volt ezekre a dolgokra, de semmit nem tudott róluk. Én se tudtam semmit, de volt egy elképzelésem róla, hogy mi lehet ez.

³ Orlandi bibliográfiájának számítógépes bölcsészettel foglalkozó része itt érhető el: <http://www.cmcl.it/~orlandi/pubinf.html>.

⁴ Koenraad DE SMEDT, Hazel GARDINER, Espen ORE, Tito ORLANDI, Harold SHORT, Jacques SOUILLOT and William VAUGHAN, eds., *Computing in humanities education: A European perspective* (Bergen: University of Bergen, 1999), <http://gandalf.aksis.uib.no/AcoHum/book/>.

⁵ Domenico FIORMONTE e Lorenzo PERILLI, a cura di, *La macchina nel tempo: Studi di informatica umanistica in onore di Tito Orlandi* (Firenze: Le lettere, 2011).

⁶ Silvio Ceccato (1914–1997) a milánói Kibernetikai Központ (Centro di Cibernetica) alapítója és igazgatója volt, egyben Európában az első, aki „az önszervezés kibernetikus elvét a fogalomalkotás és a nyelv területén alkalmazta”. Lásd: Ernst von GLASERSFELD, „Obituary: Silvio Ceccato (1914–1997)”, *Cybernetics and Systems* 29, no. 3. (1998): 213–214, <https://doi.org/10.1080/019697298125687>.

Elkezdünk kártyákat lyukasztani (nem tudom, mi lett velük), majd a projekt leállt, mert más dolgom volt. Szóval megpróbáltuk, de feladtuk.

JN: Túl bonyolult volt?

TO: Nem, más dolgok elterelték a figyelmünket! Tudja, hogy megy ez – csak egy ötlet volt. De nekiálltunk kártyákat lyukasztani, mert könnyű volt belátni, hogy ha van egy olyan rendszerünk, amely adott sorrendbe tudja rendezni és össze tudja hasonlítani a szavakat, akkor egy olyan kritikai kifejezőmód és módszerrel áll a rendelkezésünkre, amellyel el lehet kezdeni kísérletezni. Vagyis csak azt mondom, hogy olyan ötletéről volt szó, amely azokban az években még érvényesnek tűnt. Ez volt hát az első találkozásom a számítástechnikával. Persze a lehetőség élménye megmaradt a fejemben.

Az 1970-es évek végére nagy mennyiségű, kopt kéziratokkal és más irodalommal kapcsolatos információt gyűjtöttem össze. Egyre nagyobb kihívást jelentett az információ kezelése, így azt gondoltam: „meg kell próbálni automatizálni a folyamatot”. Felvettem a kapcsolatot néhány céggel, például a Nixdorf Computer AG-val⁷ és másokkal, hogy felmérjem az akkor elérhető lehetőségeket. De csak akkor indult be az egész, amikor a római egyetem Centro di calcolo (Számítóközpont) részlegéhez fordultam.

JN: Miért cégekkel próbálkozott először, ahelyett, hogy eleve a számítóközpontba ment volna?

TO: Be kell vallanom, nem voltam túl bizakodó a számítóközpont szervezetével kapcsolatban, de leginkább attól tartottam, hogy egy bölcsész tudóst nem várnának tárt karokkal. Valójában azonban rá kellett jönnöm, hogy a központ segítségemre siető munkatársai nagyon is jó szakemberek, kifejezetten elégedett voltam a közös munkával.

JN: Mások is foglalkoztak már ekkoriban számítógépes bölcsészettel az egyetemen?

TO: Igen, a már említett Ceccato professzor, de ő már arra is rájött, hogy ezek a gépek bizonyos értelemben képesek gondolkodni. Elragadó személyiség volt,

⁷ A Nixdorf Computer AG (NCAG) akkor jött létre, amikor a Labor für Impulstechnik alapítója, Heinz Nixdorf 1952-ben Essenben felvásárolta a kölni Wanderer-Werke vállalatot. Eredetileg a lyuk-kártyaszektor számára gyártott termékeket, majd az 1960-as évektől a vállalat többek között magában álló, programozható gépeket gyártott kis- és középvállalatok számára; közülük a Nixdorf 820 volt az első. Lásd: „The products of Nixdorf Computer AG”, hozzáférés: 2020.03.02, <http://www.hnf.de/en/museum/nixdorf-wegbereiter-der-dezentralen-datenverarbeitung/the-products-of-nixdorf-computer-ag.html>.

ugyanakkor kicsit hajlamos volt elrugaszkodni a valóságtól. Egyike volt azoknak az embereknek, akik olyan lelkesek, hogy alig érinti a lábuk a földet.

JN: Túl lelkes volt a technológia iránt?

TO: Igen, de ennél összetettebb a dolog. Mindennek két oldala van: jó és rossz. Úgy értem, felismerte a kapcsolatot a tiszta gondolkodás és az automatikus eljárások között. Ezt látta, de aztán a valóságtól elrugaszkodva állt hozzá.

JN: Akkor ez voltaképpen kicsit a mesterséges intelligencia története is, nem?

TO: Pontosan, de arra is megtanított, hogy óvakodjak a mesterséges intelligenciától, mert nem szeretem a bizonytalanságot. Ez a lehetőségekkel van összefüggésben, ami megint egy másik kérdés. Am valódi mesterem a számítástechnika és a számítógépes bölcsészet területén Luigi Cerofolini volt. Alkalmazott matematikusként a számok logikai elméletével és sok minden mással is foglalkozott. Rengeteget tanultam tőle arról, mit tekintünk számítógépnek és számítógépes rendszernek, és mit nem. Talán ezt nevezném az igazi fordulatnak a számítógépes bölcsészettel kapcsolatos kutatásaimban. Ő egy nagyon gyakorlatias, nagyon lényegre törő ember volt. A megközelítésében nem volt helye sarlatánságnak,⁸ és utálta a mesterséges intelligenciát.

JN: Ami a kutatásaiban bekövetkezett fordulatot illeti, kikkel dolgozott még ennek bekövetkezése előtt (leszámítva a mesterséges intelligenciával foglalkozó kollégákat)?

TO: Megpróbálom helyesen rekonstruálni az események időrendjét. Az 1950-es években sokat tanultam a mesterséges intelligenciáról ettől a zseniális embertől. Utána egy sötét időszak következett. Majd az 1970-es években a római egyetem számítóközpontjának (Centro di calcolo) a kiváló munkatársaival dolgoztam. Mirella Schaerf, az igazgató nagyon segítőkész volt. Mérnök volt, aki megértette a problémáimat, és rendelkezésemre bocsátott egy adatbázis-kezelő rendszert (ez volt az Omnidata), amely ekkor az UNIVAC nagyszámítógépén futott. Elmagyarázta, hogyan működik, és szabad hozzáférést biztosított a rendszerhez. A központ munkatársai egyes gyakorlatias ügyekben nagyon segítőkészek voltak, másokban viszont nem. A módszertani problémákat magamnak kellett megoldani.

Később találkoztam Luigi Cerofolinival, akitől megtanultam a módszertani kérdéseket is, és két olyan dolgot fedeztem fel, amelyek alapján „fordulatról” beszéllek. Először is, megismerkedtem a Turing-géppel, amelyről korábban egyáltalán

⁸ A sarlatánok kifejezésnek a számítógépes bölcsészettudomány meghatározásában Orlandinál nagy jelentősége van, a már említett, McCartyval zajlott vitájában is központi fogalom. Vö. ORLANDI, „The Scholarly Environment of Humanities Computing” (A szerk.).

nem hallottam. Másodszor pedig ekkor fedeztem fel a Unixot. Bár a számítógépes bölcsészettudomány képviselői nem akarják ezt elismerni, meggyőződésem, hogy elméleti szempontból a Turing-gép, gyakorlati szempontból pedig a Unix rendszer a legfontosabb. Cerofolini tanította meg nekem, hogy a Unix valójában nem operációs rendszer, hanem környezetrendszer. Csupán ennyi az az alap, amelyre a munkánk épül. Azt hiszem, ez egy nagyon fontos háttér, és ebben rejlett a fordulat.

JN: Tudom, hogy publikált már a témában,⁹ de kifejténé bővebben az univerzális Turing-gép elméletét, és azt, hogy milyen hatással volt ez a munkájára?

TO: Nemcsak a munkámra, az életemre is hatással volt! Komolyra fordítva a szót, megértettem, hogy a Turing-gép titokzatos és egyfajta misztikus-filozofikus kapcsot képez a logika, a gondolkodás és valami anyagi értelemben történő dolog között, legyen az számítógép vagy bármi más. A számítógép ugyanis több, mint amit általában „számítógépnek” nevezünk. Bármi, ami diszkrét mennyiségeken automatikus eljárásokat képes végrehajtani, számítógép. Épp ezért lehet az elméleti ötleteinket tettekre fordítható módon kifejezni. Ez az, amit igazán szeretek a számítógépes bölcsészettudományban, mert a bölcsészettudományban bármit meg lehet vitatni, és mindenkinek igaza van. Hogy lehet ellenőrizni, hogy Vergilius jó költő vagy nem? Hogy lehet ellenőrizni, hogy az adott verset tényleg Vergilius írta-e vagy sem? stb. Ezek évszázadok óta tárgyalt kérdések, és mindenkinek igaza van! Nos, én nem akarom azt mondani, hogy a bölcsészet területén minden megoldható számítógépekkel. De egyes kérdésekre legalább annyit mondhatunk, hogy „hic Rhodus, hic salta”.¹⁰ Ha van egy ötleted, képletté alakítod, és beleteszed egy anyagi értelemben valóságos valamibe. Ezt Turing előtt nem lehetett megcsinálni. Turing után viszont igen; ezért olyan csodálatos a Turing-gép.

JN: Ön szerint ez a hozzájárulás meghatározza a számítógépes bölcsészeti tevékenységet?

TO: Tudja, a bölcsészettudományi kutatásokat én két területre szoktam osztani. Az egyik területen a logika az irányadó, a másikon pedig az a valami, amit intuíciónak nevezhetnénk. Az intuíció nem szabályozható, ezzel a képességgel vagy

⁹ Lásd például: Tito ORLANDI, *Ideas for a theoretical foundation of humanities computing* (London Seminar, King's College London: Unpublished [2000]), <http://www.cmcl.it/~orlandi/pubbli/saggi-omio.pdf>; Tito ORLANDI, „Is humanities computing a discipline?”, in *Jahrbuch Für Computerphilologie*, Hg. Georg BRAUNGART, Karl EIBL und Fotis JANNIDIS, 4, 51–58 (Paderborn: Mentis, 2002), <http://computerphilologie.uni-muenchen.de/jg02/orlandi.html>.

¹⁰ „Itt van Rhodosz, itt ugorj”, értsd: „Most mutasd meg, mit tudsz”. Ered. Aesopus 33. (Perry-index), a dicsekvő atlétáról szóló fabulája. (A szerk.)

rendelkezünk vagy nem. Az intuíció nagyon nehezen vizsgálható: egyikőnk azt mondja valamire, hogy „jó megérzés”, ami a másik szerint „rossz megérzés”.

Ám ha valaki előáll egy történeti javaslattal, akkor azt logikusan kell felépíteni. Ha nincs jól felépítve logikailag, ellentmondásos lesz, és nem állja meg helyét. A bölcsészettudománynak ez a része automatizálható, aminek óriási hatása lehet, mert a számítógépek (banális, de ezzel kezdődik) olyan mennyiségű emléket képesek kezelni, amelyet az emberi agy nem. Ha pedig a módszerünket (vagy logikánkat) óriási mennyiségű anyagra tudjuk alkalmazni, akkor konkrétan látni fogjuk, milyen esetekben nem működik.

JN: Amikor Milánót említette, azon gondolkodtam, vajon találkozott-e Busával és az ő munkájával?

TO: Busával viszonylag későn, és nem Milánóban találkoztam. Hogy valóban Busa atya lenne-e a számítógépes bölcsészet ősforrása, az kényes kérdés. Ebben az ügyben két megfigyelésre korlátozom magam: először is, bár kapcsolata az IBM-mel (amely ekkoriban még nem foglalkozott valódi számítási feladatokkal, mint az UNIVAC) természetesen bizonyított, arra vonatkozóan azonban Busa visszaemlékezésein kívül egyáltalán nincs bizonyítékunk, hogy valódi számítási feladatokban gondolkodott. Másodsorban pedig Busa nem rendelkezett hagyományos értelemben vett nyelvészeti vagy szemiotikai háttérrel, és a munkássága távol esett az ekkoriban zajló számítástechnikai fejlesztésektől.

A számítógépes bölcsészet valódi kezdeteinek alighanem konkrét kísérleteket kell tekintenünk, például a mesterséges fordítás vagy az automatikus fordítás területéről, amely ugyan nem volt hibáktól mentes, de most nem ez a fontos. Itt természetesen olyanok munkásságára utalok, mint William Weaver, Norbert Wiener és mások. Figyelembe kell venni ugyanakkor a régészet bizonyos területeit is, különösen Jean-Claude Gardin kísérleteit és az új régészetet, Amerikában és máshol. Ott igazán érdekes dolgokat találni. Hibák persze történtek, kicsit próbaszerencse alapon működött a dolog. De nem értek egyet azzal, hogy Busa atyát is az úttörők között lehetne említeni. A manapság általa elfoglalt pozíció nem csupán téves, de félrevezető is.

JN: A Busával kapcsolatos kutatásaimban jelenleg az a munkahipotézisem, hogy Busa legendája egy bizonyos mértékig hasznos fikció.

TO: Pontosan!

JN: És ezzel nem szeretném megkérdőjelezni a munkásságát, de úgy gondolom, hogy a legendáját a közösség egyfajta alapítási mítoszként kapta fel. Olyan alak lett, akire ki lehet vetíteni, köré lehet szervezni a dolgokat. Úgy gondolom azonban, maximálisan egyetértve az Ön imént kifejtett álláspontjával, hogy valódi

intellektuális kapocs a korai írásaiban nagyon nehezen fedezhető fel (persze ez a kutatás későbbi fázisában még bizonyulhat tévesnek). Mindenesetre az egyik dolog, amit szeretnék megérteni, épp az, hogy hogyan alakult ki ez a „hasznos fikció”. Milyen formálódási és hagyományozódási folyamatok alakították ki?

TO: Ha szabad ezt mondani, Busa nagyon hatékonyan tudta eladni magát. Tudjuk, ez a mai világban mennyire fontos. Azt is el kell ismernünk, hogy megvolt hozzá a képessége, hogy megértse, általában miként képzelik el az emberek a számítógépek humán tudományokon belüli alkalmazását. Hihetetlenül intelligens ember volt, ez vitathatatlan. De nem tudta felismerni azokat a változásokat – és ez összefüggésben állhat azzal a fontos ténnyel, hogy jezsuita volt –, amelyeket a nyelvészeti és a gondolati szemléletben a Turing-gép okozott. Meggyőződésem szerint, még ha talán tévedek is, Busa atya és Turing két külön dolog, köszönőviszonyban sincsenek egymással. Mindezek alapján belátható, miért nem értek egyet azzal, hogy ő is úttörő lett volna. A régi olasz irodalmi szövegeken végzett sokkal komolyabb munka, amely Mario Alinei és D’Arco Silvio Avalle nevéhez fűződik, nem Busa csoportjából származik.

JN: Megkérhetem, hogy nevezzen meg néhány olyan további projektet, amelyek véleménye szerint fontos előrelépést jelentettek?

TO: Az első kritikai kiadás tekintetében például Peter Robinson Chaucer-kiadása¹¹ jut eszembe. De az ilyen projektek általában nyitottak. Fenntartom véleményemet, hogy nem egy adott projekt teljesítése a fontos, hanem az a módszertani szemlélet, amelyből ered. Ebben a tekintetben Robinson Chaucer-kiadása nagyon érdekes volt. Kiegészíteném még ezt Jean-Claude Gardin munkásságával (lásd alább).

JN: Végzett valaha hivatalos számítástechnikai képzést?

TO: Egyáltalán nem. Tankönyvekből képeztem magam. Az egyetem számítógépközpontjában dolgozó kollégák időnként gyakorlati útmutatást nyújtottak, és számos alkalommal találkoztam Luigi Cerofolinivel. Barátok lettünk, idővel nagyon jó barátok. Emlékszem, hogy 1980-ban vagy 1981-ben Amerikába mentem, és körbenéztem az egyetemi könyvesboltokban, hogy milyen számítástechnikával és kapcsolódó területekkel foglalkozó könyveket árulnak. Megszereztem és a mai napig őrzöm a számítástechnika elméletével és tudományával kapcsolatos alapműveket, amelyeket a számítógépes bölcsészek ma egyáltalán nem emlegetnek, és furcsa, ahogy erről beszélnek.

¹¹ Peter ROBINSON, ed., *Geoffrey Chaucer, the wife of Bath’s prologue*, CD-ROM (Cambridge–New York: Cambridge University Press, 1996).

Én mindig tanultam, mindig próbáltam egyre mélyebbre ásni. Nem győzöm eleget szajkózni, hogy a legtöbb, számítógépes bölcsészettel foglalkozó kollégám nem olvas az úgynevezett „témájáról”. Mennyit tud egy szakmunkás, egy számítógépes bölcsészettel foglalkozó szellemi szakmunkás a számítástechnika irodalmáról? Amennyire én látom, jellemzően semmit vagy majdnem semmit. Ez a legnagyobb hiányosság a számítógépes bölcsészet területén; persze Önök az UCL-nél (University College London – a szerk.), illetve a King’s College London is kivételnek számít. Általában véve azonban az ilyen ügyeknek nincs helye a diszciplínán belül, és erre persze mindannyian panaszkodunk. De hát megvan-nak a magunk hibái!

Én például összeállítottam egy olyan könyvtárat, amely a nyelvésztől és a kódolási elmélettől Jean-Claude Gardin *Archaeological Constructs: An Aspect of Theoretical Archaeology* című könyvéig¹² terjedt. Ha könyvtárunkban Gardin műveitől a Turing-gépet tárgyaló művekig szerepelnek kötetek, könnyebben megláthatjuk a közöttük fennálló kapcsolatokat.

Véleményem szerint még mindig a számítógépes bölcsészet korai szakaszában vagyunk – a módszertani kidolgozottságot tekintve. Az emberek forradalomról és radikális változásról beszélnek, amelyek valójában a dolgok felszínén zajlanak. Az igazán mély változásokhoz azonban nem évtizedek, hanem évszázadok kellenek. Tudja, a számítógépes bölcsészet mondjuk az 1940-es évek végétől létezik. Majdnem százéves. Ha körbekérdezzük az embereket (mármint a mi körünkön kívül), és megkérdezzük, szerintük „mi az a számítógépes bölcsészet?”, azt fogják válaszolni: „könyvtárak, katalógusok” vagy „szöveggyűjtemények”. Na de mi a számítógépes bölcsészet abban, hogy ennek vagy annak a szövegnek a PDF-változata elérhető? Ez nem számítógépes bölcsészet!

JN: Egyáltalán nem. Sajnos sok ember úgy gondolja, hogy ennyiben merül ki. Benyomásom szerint ez különösen azóta jellemző, amióta átálltunk a digitális bölcsészet (*DH, Digital Humanities*) kifejezésre, amelyet manapság gyakran használnak a számítógépes bölcsészet (*Humanities Computing*) helyett.

TO: Igen, sajnos az emberek nem tudják, mi az, hogy digitális! Amikor azt mondják, digitális, arra gondolnak, hogy „elektronikus”. De mit tudnak a bölcsészek például a digitális és az analóg közötti különbségről? Azt gondolják, ez egy magától értetődő fogalom, amelynek nem kell, hogy tudatában legyenek. Én mindig is a matematika ellen voltam, abban az értelemben, hogy amire matematikaként gondolunk, az valójában a környezetünk, ezért nem akarom matematikának hívni. Ez a logika. És a „digitális” nem feltétlen „elektronikus”, egyáltalán nem.

¹² Jean-Claude GARDIN, *Archaeological constructs: An aspect of theoretical archaeology* (Cambridge–New York: Cambridge University Press, 1980).

JN: Említette a „forradalom” szót, amely kirívóan gyakran előkerül a digitális bölcsészettel és számítógépes bölcsészettel foglalkozók körében. Ez a szó szerintem több szempontból is zavarba ejtő, mert – legalábbis számomra – a forradalom azt jelenti, hogy ledöntjük a korábbi, romlásnak indult múltat, és elkezdünk egy dicső jövőn dolgozni. Szóval talán nem feltétlenül ebben az értelemben használják a szót – tudna kicsit arról beszélni, milyen kontextusban találkozott a kifejezéssel, és hogy Ön szerint mit értenek e kifejezés alatt a számítógépes bölcsészet területén?

TO: Nem tudok – abban az értelemben, hogy ha ilyen jelenségeket szeretnék leírni, akkor csak ezt a szót találom. A számítógépek ma már nagyon nem azok a „furcsa gépek”, amelyek régebben voltak. A bölcsészettudomány jeles területein kevesen képzeltek volna, hogy a számítógépek és a bölcsészet útjai keresztezhetik egymást. Tudja jól, hogy azt mondták, „a számítógépek matematikára valók, a bölcsészet a gondolkodásra. A számítógépeknek semmi közük a nyelvekhez vagy a történeti eseményekhez.” Azok, akik felismerték a két terület összekapcsolásának lehetőségét, forradalomról beszéltek, de nem olyan értelemben, hogy a régebbi dolgokat megvetették volna. Inkább arról volt szó, hogy várakozásaik szerint az új eszköz alapjaiban felforgatta volna a bölcsészettudományok korábbi szokásait. Olyan értelemben forradalom ez, mint ahogy Elizabeth Eisenstein¹³ beszél nyomtatási forradalomról. Semmi bajunk a kéziratokkal, de a nyomtatás forradalmat hoz magával.

JN: De vajon miért használják a kifejezést ma is? Szerintem ma már nem igazán releváns. Meghatározó kifejezés maradhat, de ma már olyan körülményekre utal, amelyek elmúltak.

TO: Sajnos a „forradalom” kifejezést régóta azért használják, mert egyfajta, úgy mondanám, „szociológiai megalapozottságra” tett szert. Az emberek ma azt látják, hogy megváltoztak az olvasók, megváltoztak a könyvtárak, megváltoztak a régészeti ásatások és így tovább. De csak azért, mert megváltoztak a szerepek és az eszközök – egy ásatás során ma már telekamerákat és egyéb nagyszerű eszközöket használunk, és azonnal megkapjuk a mérési eredményeket –, már mindjárt azt mondják: „Ez aztán a forradalom!” Ez egyáltalán nem forradalom – inkább ahhoz hasonlít, hogy ma már bárki számára elérhető a mikrohullámú sütő, míg ötven évvel ezelőtt nem. Ez lenne a forradalom? „Hát persze, hogy az, még hozzá értelmes forradalom.” Az étel azonban ebből a szempontból többé-kevésbé ugyanaz maradt.

¹³ Elizabeth EISENSTEIN, *The printing press as an agent of change: Communications and cultural transformations in early-modern Europe* (Cambridge–New York: Cambridge University Press, 1980), <https://doi.org/10.1017/CBO9781107049963>.

JN: Mikor találkozott először a számítógépes bölcsészek közösségével?

TO: Abban a kivételes helyzetben vagyok, hogy azt mondhatom, nem találkoztam a közösséggel, hanem láttam, amint létrejön. Már a kezdetek kezdetén megismerkedtem például Antonio Zampollival. Pisában dolgozott, a CNR Istituto di Linguistica Computazionale (Számítógépes Nyelvészeti Intézet) egységénél, így láthattam, hogyan alakulnak ott a dolgok. Willard McCartyt itt, Rómában ismertem meg egy olyan találkozón, amelyet a kanadai nagykövetség szervezett, amikor ő még Kanadában dolgozott. Így már a karrierje legelején megismertük egymást. Azt hiszem, nagyon sok olyan emberrel találkoztam, akik már a kezdeteknél is jelen voltak.

Azt kell mondanom, Gardin egy kivételes eset, mert ő valóban ott volt a számítógépes bölcsészet ősforrásánál. Gardin azok közé tartozott (nemrég hunyt el), akik iránt a legmélyebb tiszteletet érzem. Már az 1950-es években reflektált a számítástechnika lehetőségeire, csak éppen senki sem tudott róla. Nagyon tartózkodó ember volt, így a történetét sem nagyon ismerik. Egészen addig én sem tudtam a munkásságáról, amíg nem olvastam a régészeti konstrukciókról szóló könyvét. Nem a kezdetek kezdetén találkoztam vele, hanem akkoriban, amikor a számítógépes bölcsészet Franciaországban terjedni kezdett. Németországban találkoztam Manfred Thallerrel (lásd: 13. fejezet),¹⁴ aki szintén alapos ismeretekkel rendelkezik. Persze nem lett szerencséje csak azért, mert jó elméleti szakember, de ez előfordul.

JN: Mit ért az alatt, hogy Thaller nem volt szerencsés?

TO: Ahhoz, hogy professzor legyen, Thallernek kompromisszumot kellett kötnie, „venire a patti”.¹⁵ Ő állította össze a kölni kéziratok reprodukcióinak lenyűgöző gyűjteményét (lásd: 13. fejezet). Nagyon szeretem ezt a gyűjteményt. Számítógépes bölcsészetnek számít? Nem. Vagy igen, de csak komoly fenntartásokkal.

JN: Ez bizonyos szempontból kapcsolódik ahhoz a korábbi kérdésemhez, hogy meglátása szerint mennyire értékelték, és hogyan reagáltak a számítógépes bölcsészet területére azok, akik nem foglalkoztak a témával?

TO: Szkeptikusan, vagy inkább a *negazione* (tagadás) és a *rifiuto* (elutasítás) közötti skála valamely fokával. Akár azt is mondhatnám, „jogosan”, mert a számítógépes bölcsészet területén zajló vállalkozások általában nem voltak elég megalapo-

¹⁴ A 13. fejezet: „It’s Probably the only Modestly Widely Used System with a Command Language in Latin: Manfred Thaller and Julianne Nyhan (interview)”, in Julianne NYHAN and Andrew FLINN, *Computation and the Humanities: Towards an Oral History of Digital Humanities*, 195–208 (Cham: Springer, 2016), <https://doi.org/10.1007/978-3-319-20170-2>. (A szerk.)

¹⁵ Az eredeti szövegben: „To become a professor Thaller had to »venire a patti«, to compromise.” 83.

zottak ahhoz, hogy felkeltsék a számítástechnikával nem foglalkozó bölcsészek figyelmét. Természetesen a konkordanciák és hasonló dolgok előállítását ugyanazek a bölcsészek azonnal jóváhagyták. Az ilyen eszközök nagyon fontosak voltak, de semmi elméleti nincs bennük. Az Oxford Text Archive¹⁶ csodálatos dolog, utána pedig jött a Google. Belátható, hogy ha valami gyakorlati szempontból hasznos, akkor azt értékelik, ám az ilyen példák semmilyen módon nem kapcsolódnak az egyes bölcsészkatatók módszertanához és tanulmányaihoz.

Nagyon fontosak az elméleti nyelvészet azon fejleményei, amelyek a mesterséges fordítással kapcsolatos szerencsétlen kísérletekből erednek.¹⁷ Ebben az ügyben nem csak az itt, Olaszországban zajló kísérletekre gondolok. Geoffrey Sampson és a British National Corpus¹⁸ és sok hasonló vállalkozás azért is jelentősek, mert kéz a kézben jártak Sampson két témában megfogalmazott nézeteivel. Az első a szintaktikai nyelvészet és az a vitairat,¹⁹ amelyet Chomsky ellenében írt. (Aki viszont kiváló példa a számítástechnikai alapelvek és a nyelvek közötti tényleges együttműködésre. Sajnos fenn kell tartanom, hogy a filozófiai alapjai nem voltak elég szilárdak, és úgy tekintett a történeti nyelvekre, mint amiket a természet adott. Ennek ellenére is nagyon érdekes figura ebből a szempontból.) Az a könyv csodálatos. A második a kódolási elveket érinti, amely összekapcsolódott az ábécé-elmélettel. Sampson írt egy kiváló könyvet az ábécéről:²⁰ különös ez a British National Corpus létrehozójától, de éppen erről beszélek. A bölcsészettudomány úgy fejlődhet, ha reflektálunk a számítógépes alkalmazásokra, és ez a könyv nagyszerű példa erre.

JN: Ami az el nem ismert projekteket illeti, meglátása szerint ezeket azért nem fogadták jól, mert az alkalmazhatóságuk nem volt világos a bölcsészettudományok számára?

TO: Szerintem ezeket egy olyan jelenség miatt nem fogadták el, amelyről őszinte sajnálattal kell nyilatkoznom. Amikor valaki komolyan elkezd dolgozni az automatizáláson (vagyis számítógépeken, de az automatizálás értelmében), akkor szinte azonnal jön a felismerés, hogy a bölcsészettudományokban nagyon

¹⁶ „A University of Oxford Text Archive a felsőoktatásban, kutatásban, tanításban és tanulásban használt elektronikus irodalmi és nyelvészeti forrásokat fejleszt, gyűjt, katalogizál és tárol.” Lásd: <http://ota.ox.ac.uk/>.

¹⁷ W. John HUTCHINS, ed., *Early years in machine translation: Memoirs and biographies of pioneers*, Studies in the history of the language sciences (Amsterdam: John Benjamins Publishing Company, 2000), 97.

¹⁸ „A British National Corpus (BNC) 100 millió szóból álló, változatos forrásokból származó írott és beszélt nyelvi mintákat tartalmazó gyűjtemény, melynek célja, hogy széles keresztmetszetet nyújtson a beszélt és írott brit angol nyelvről, a késő huszadik századtól kezdve.” Lásd: <http://www.natcorp.ox.ac.uk/>.

¹⁹ Geoffrey SAMPSON, *The “Language Instinct” debate: Revised edition* (London–New York: Continuum-3PL, 2005).

²⁰ Geoffrey SAMPSON, *Writing systems: Revised edition* (Sheffield–Bristol: Equinox Publishing Ltd., 2015).

gyakran nem tudjuk pontosan, mit csinálunk, és ezt nagyon nehéz elfogadni. Ez az én személyes tapasztalatom: hogy mondhatnám azt egy, például olasz irodalommal foglalkozó kollégámnak, hogy nem ismeri rendesen az általa tanulmányozott szövegeket? Hogyan mondhatnám el neki, hogy nem érti, mi az a szöveg, és hogy milyen alapon mondhatjuk azt, hogy ezt a szöveget Dante írta vagy nem? Vagy még inkább, mi a jelentése a helyesírásnak ebben vagy abban a kéziratban? Mi a különbség egy levél tényleges anyagi kialakítása, illetve a nyelv írott és beszélt apparátusának részeként értelmezett graféma között?

A kollégám azt mondaná, hogy megőrültem, és ezek nem problémák, vagy ha igen, akkor nem szeretné tanulmányozni őket. Ennek valószínűleg az az oka, hogy intuitív problémákról van szó. Ha azonban meg kell tanítani egy gépet arra, hogyan kezelje az ilyen adatokat, akkor a gépnek mindenről pontosan meg kell mondani, hogy micsoda, és akkor jövünk rá, hogy nem is tudjuk a válaszokat a fenti kérdéseimre! A legtöbb bölcsészt épp ez tántorítja el, mert képtelen mindezt elfogadni. Ez egy hosszú folyamat, de idővel az átlagos bölcsész kutató talán elfogadja, hogy az ilyen jelenségek formális definíciójával kapcsolatban nehézségek merülhetnek fel. És ez „la scommessa”, fogadás a jövőre, mert az egyszer biztos, hogy a jövőben minden számítógépesítve lesz. Nem kérdés, hogy akarjuk-e ezt vagy sem – meg fog történni. És ha ez a helyzet, akkor a mostani generációkon múlik, hogy az egyes tudományágakhoz tartozó adatok milyen pontosan lesznek számítógépesítve. Emiatt, ha nem is vagyok megszállottja a kérdésnek, de legalábbis próbálok minél többet gondolkodni erről a lényeges problémáról.

Mint tudjuk, ma az infrastruktúrák jelentik a legújabb trendet. Az Európai Közösség számos területen javasolja infrastruktúrák kialakítását, de mindenki a maga útját járja. Hogy hangzik az az ötlet, hogy rendezzünk találkozókat, ahol megvitatjuk, hogyan kellene megszervezni a számítógépes bölcsészet infrastruktúráit? Ez óhatatlanul költséggel jár, de az infrastruktúrákat akkor is ki kell alakítani, európai pénzzel vagy anélkül. Az egyetemeknek előbb-utóbb ki kell alakítania ilyeneket, és ez egyre kevesebbe fog kerülni. Kutatási problémát nem látok az infrastruktúrák területén, sőt, épp ellenkezőleg. Természetesen hatalmas mennyiségű adatunk lesz, és ennek csak örülhetünk. De hogy kerül ez az adat digitális formátumba? Ehhez nagyon kifinomult szemlélet kell, és ezt kevesen fogják megérteni. Willard McCarty, Geoffrey Rockwell és Manfred Thaller, egy, két, három ember érteni fogja, a jelenség viszont közben elterjed a világon. De természetesen nem lehetünk pesszimisták! Azt azonban észre kell vennünk, hogy ez a következő évek nagy kihívása. Próbáljuk erről meggyőzni a felelős személyeket, még ha ez nem is könnyű.

JN: Számos interjúalany idézte fel, hogy amikor elmentek az első, számítógépes bölcsészettel foglalkozó konferenciájukra, azt tapasztalták, hogy az emberek nagyon barátságosak, és időnként úgy érezték, sokkal barátságosabb a közös-

ség, mint a saját bölcsészettudományi tudományágukon belül. Azt szeretném megkérdezni, mi a véleménye erről?

TO: Igen, egyetértek. Nagyon nagy volt a „cameratismo” (bajtársiasság). Egyszerűen ez volt a szokás, mint ahogy a tágabb számítástechnikai közegben is. A bölcsészettudományokban azonban más volt a helyzet. Amikor például újtára indult az Oxford Patristic Conference,²¹ egyáltalán nem volt olyan nagy, mint ma, egy tárgyalóteremben zajlott. Akárhogy is, mindenki más iskolába tartozik. Úgy gondolom, mára a bölcsészettudományokban is megváltozott a helyzet, ott is kialakult az a bajtársias érzés, amely korábban hiányzott.

JN: Említett néhány embert, aki különleges hatással volt Önre. Van még valaki, akit esetleg zárásképpen megemlítené?

TO: Nem, attól tartok, ilyen helyzetekben óhatatlanul kifelejtünk valakit, de legtöbbjüket már megemlítettem.

(„*hic Rhodus, hic salta: Tito Orlandi and Julianne Nyhan*”. In Julianne NYHAN and Andrew FLINN, *Computation and the Humanities: Towards an Oral History of Digital Humanities*, 75–86. Cham: Springer, 2016. <https://doi.org/10.1007/978-3-319-20170-2>.)

Fordította: *Maczelka Csaba*

²¹ Egy 1951-ben alapított, a patrisztika területén az egyik legrangosabbnak számító konferenciasorozat. Lásd: <http://www.oxfordpatristics.com/#!conferenc/c1jxp>.

MŰHELY

KALCSÓ GYULA

Kézírtos források tudományos elektronikus kiadása

Tanulmányom a kézírtos források tudományos elektronikus kiadásáról, elsősorban a kiadás megtervezésének, valamint szabványos kódolásának kérdéseiről szól, nem foglalkozik viszont a kiadás publikálásának a problematikájával. A tudományos elektronikus kiadás fogalmát Maróthy Szilvia munkái alapján használom, a szerző az angol *digital scholarly edition* kifejezés magyar megfelelőjeként javasolja több, a közelmúltban megjelent cikkében.¹ Az ilyen kiadást Patrick Sahle definíciója alapján kritikai reprezentációnak tartja, amely „létező forrásokat, történeti dokumentumokat dolgoz fel, forrást reprezentál és azt metaadatokkal látja el, valamint szövegkritikai munkát végez a forrás reprezentálásakor”.² Sahle nyomán emlékeztet azon elterjedt nézetre, hogy a digitális kiadás paradigmaváltást jelent:

Míg a hagyományos kiadásokban a szöveggondozó írta a szerkesztett szöveget, a digitális kiadásban az a források feldolgozási lépéseiből épül fel, a képi reprezentációtól az átíráson át a kritikai apparátusig. A források különféle reprezentációinak (és n. b. a szöveg különféle forrásainak) „elrejtésére” ezen túl nincs szükség, azok teljes terjedelmükben részei lehetnek a kiadásnak. E változásnak a következménye Sahle szerint a szerkesztés során a szöveg relativizálódása és multiplicitásának (összetettségének és többszörösségének) előtérbe helyeződése.³

Debreczeni Attila *Az elektronikus kiadás színe és visszája* című írásában felhívja a figyelmet arra, hogy „a digitális tudományos kiadásnak sem megvalósulásai nincsenek nagy számban, sem konszenzuális teóriája, minden számottevő eredmény ellenére sem”.⁴ Mindezek miatt indokolt lehet olyan teoretikus megfontolások

¹ MARÓTHY Szilvia, „Elektronikus szövegkiadások a könyvtárban”, *Tudományos és Műszaki Tájékoztatás* 64, 6. sz. (2017): 298–309; MARÓTHY Szilvia, „A tudományos elektronikus szövegkiadások feldolgozása”, in *Valóságos könyvtár – könyvtári valóság: Könyvtár- és információtudományi tanulmányok 2018*, szerk. KISZL Péter és CSÍK Tibor (Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar Könyvtár- és Információtudományi Intézet, 2018), 351–356; MARÓTHY Szilvia, „Tudományos szövegkiadások a hálózaton”, *Irodalomtörténeti Közlemények* 122, 5. sz. (2018): 617–633.

² MARÓTHY, „A tudományos elektronikus szövegkiadások feldolgozása”, 351–352.

³ MARÓTHY, „Elektronikus szövegkiadások a könyvtárban”, 299.

⁴ DEBRECZENI Attila, „Az elektronikus kiadás színe és visszája”, in *A humán tudományok és a gépi intelligencia*, szerk. TOLCSVAI NAGY Gábor, *A humán tudományok alapkérdései* 3, 48–61 (Budapest: Gondolat Kiadó, 2018), 49.

publikálása, amelyek középkori kéziratos források tudományos elektronikus kiadásának előkészítése, valamint a megvalósítás kezdeti, bátortalan lépései közben körvonalazódtak.⁵

A kéziratok digitális publikálása már hazánkban is elég gyakori, a középkori kódexektől a különféle levéltári forrásokon át az írói levelezésig mindenféle kéziratos műfaj képviselteti magát. Azt azonban aligha mondhatjuk, hogy valamiféle egységes módszertant tükrözne a forrásdigitalizálási és -közlési gyakorlat. Tószegi Zsuzsanna *A szövegdigitalizálás döntési folyamata* című cikkében a digitalizálásnak három szintjét különbözteti meg: a reprodukzív, a reprezentatív és az interpretatív szintet.⁶

A reprodukzív szint a forrásmű formai és tartalmi jegyeit egyaránt tükrözteti (az esetleges hibákkal, eltérésekkel együtt). A digitalizált változat az eredeti művel gyakorlatilag egyenértékű, azzal egyező hatást vált ki. Ebbe a csoportba elsősorban a faksimile állományok (képfájlok) tartoznak

– írja Tószegi.⁷ Az ilyen közlések a forrás különböző minőségű képét publikálják többnyire valamilyen szabványos formátumban, esetenként a képet tartalmazó PDF-ben. Ilyen például az Országos Széchényi Könyvtár nyelvemlékes honlapja (például a *Szabács viadala* PDF-változata)⁸ vagy a Magyar Elektronikus Könyvtár (MEK) digitalizált nyelvemlékei (például a *Régi magyar nyelvemlékek* 33. darabjaként megjelent Apor-kódex DVD-mellékletének böngészhető változata).⁹ Ezek a „digitális faksimilék” jobbára az autopszia helyettesítésére valók, nem teszik azonban lehetővé a keresést a szövegekben.

A reprezentatív szint a forrásmű tartalmát helyezi előtérbe, de alapvetően nem változtatja meg a szöveg lineáris olvasatát. Ezen a szinten az analóg szövegből digitalizált szöveget állítunk elő, amelynek információtartalma a számítógép nyújtotta szokásos eszközökkel könnyebben kereshető.¹⁰

⁵ A munkálatok a 2018 őszén, az Eszterházy Károly Egyetemen megalakult Digitális Bölcsészeti Kutatócsoportban (<https://uni-eszterhazy.hu/hu/bmk/bmk/a-bmk-hirei/c/digitalis-bolcseszeti-kutato-csoport>) indultak meg, kezdeti célként a csoport három ómagyar kéziratos forrás, a Tihanyi alapítólevél, a Halotti beszéd és könyörgés, valamint az Ómagyar Mária-siralom tudományos elektronikus kiadásának az elkészítését tűzte ki. A tanulmány a koncepció ez idáig körvonalazódott elemeit adja elő.

⁶ TÓSZEGI Zsuzsanna, „A szövegdigitalizálás döntési folyamata”, *Könyvtári Figyelő* 52, 2. sz. (2006): 245–260, 245.

⁷ Uo.

⁸ A *Szabács viadala* faksimiléje, hozzáférés: 2020.02.29, http://nyelvemlekek.oszk.hu/sites/nyelvemlekek.oszk.hu/files/mny2_b.pdf.

⁹ Az Apor-kódex faksimiléje, közzéteszi HAADER Lea, KOCSIS Réka, KOROMPAY Klára és SZENTGYÖRGYI Rudolf, hozzáférés: 2020.02.29, <http://mek.oszk.hu/16300/16329/html>.

¹⁰ TÓSZEGI, „A szövegdigitalizálás döntési folyamata”, 245.

A szövegdigitalizálás „hőskorából” származó szövegeken túl¹¹ – amelyek gyakran mutattak fel kezdetleges, mondhatni „barbár” megoldásokat¹² – mára a Tószegi cikkének megírása idején még nem létező kétrétegű PDF-ek is ilyen jellegű közléseknek számítanak, igaz, a reprodukív és a reprezentatív szint sajátos keverékének tekinthetők. Ezekben a forrás képe „mögött” megtalálható a – többnyire optikai karakterfelismeréssel (angol betűszavával OCR-rel) előállított – szöveg, amelyben ily módon kereshetünk is. Az OCR minősége – amely a történeti szövegek esetében általában nem jó –, illetőleg az utómunkára, főként a kézi korrektúrára fordított idő meghatározza a használhatóságot. Ilyen közlés például a Cornides-kódex MEK-beli változata, amely a *Codices Hungarici* sorozat nyomtatásban megjelent VI. kötetének OCR-ezett PDF-je.¹³

A digitalizálás legmagasabb szintjét a Tószegi által „interpretatív szintű” szövegközléseknek nevezett kiadások jelentik:

Az interpretatív szinten az eredeti forrás tartalmához hozzáadódik a feldolgozást végző szakemberek tudása és tapasztalata, melynek eredményeként új minőség jön létre. Az eredeti művet kiegészítő elemek (amelyek lehetnek magyarázatok, mutatók, hipertext hivatkozások, vagy a szövegtől eltérő műfajú elemek: hang- és videofájlok stb.) megbontják az eredeti szöveg lineáris egységét.¹⁴

Ezekben a kiadásokban a szerkesztők a digitális szövegben különböző hozzáadott információkat kódolhatnak, amelyek az adatbázisba szervezéshez vagy az összetettebb keresések végrehajtásához szolgálhatnak alapul. Az annotáció tartalma általában attól függ, hogy milyen célból publikálják az elektronikus kiadást. A magyar gyakorlatban van példa papíralapú kritikai kiadásokhoz hasonló apparátus használatára, de nyelvészeti annotációt tartalmazó digitális korpuszokat is fejlesztettek már. Az előbbire példák a Sermones Compilati Kutatócsoport elektronikus szövegközlései, többek között az Érdy-kódex digitális változata,¹⁵ amelyben hipertextes lábjegyzetek formájában szerepelnek a szöveggkritikai annotációk, az utóbbira pedig az MTA Nyelvtudományi Intézetében fejlesztett ómagyar korpusz¹⁶ vagy a *Történeti magánéleti korpusz*.¹⁷

¹¹ Ilyenek pl. a MEK magyar irodalmi klasszikusokat népszerű, olcsó kiadásokból közreadó digitalizált szövegei.

¹² Vö. *Filológia és digitális barbárság: A BIÖP tanulmányi napja*. ELTE BTK, 2004. március 4, hozzáférés: 2019.12.19, <http://magyar-irodalom.elte.hu/biop/barbar/>.

¹³ A Cornides-kódex kétrétegű PDF-ben, hozzáférés: 2020.02.29, <http://mek.oszk.hu/08900/08981/pdf>.

¹⁴ TÓSZEGI, „A szövegdigitalizálás döntési folyamata”, 245.

¹⁵ Az Érdy-kódex elektronikus kritikai kiadása, hozzáférés: 2019.12.19, <http://sermones.elte.hu/erdy>.

¹⁶ A Magyar Generatív Történeti Szintaxis projekt ómagyar korpusza, hozzáférés: 2019.12.19, <http://omagyarkorpusz.nytud.hu/>.

¹⁷ Az MTA NYTI Történeti magánéleti korpusza, hozzáférés: 2019.12.19, <http://tmk.nytud.hu/>.

Az „interpretatív szint” definíciója azonban kiegészítésre szorul. Tószegi elfeledkezett arról, hogy interpretálni csak már valamilyen formában létrejött digitális objektumot lehet, vagyis az interpretatív szint szükségképpen összekapcsolódik a reprodukívval, a reprezentatívval vagy mindkettővel. Az annotált képek (reproduktív + interpretatív) és az annotált szövegek (reprezentatív + interpretatív) hazánkban sem példa nélküliek. Azonban a minden kétséget kizáróan legfejlettebb elektronikus kiadásnak számító, a képet és az annotált szöveget összekapcsoló, rugalmasan kereshető, az adatokkal műveletek végzését lehetővé tevő, szabványos és időtálló formátumban létrehozott közléseket hiába keresünk.¹⁸

Mi lehet az a szabványos és időtálló formátum, illetőleg módszer a digitális filológiában, amely a magas minőségű elektronikus tudományos kiadások esetében a bölcsek követ jelentheti? A kérdés napjainkig nyitott. Aktualitását mi sem jelzi jobban, mint a Text Encoding Initiative 2019 őszén, Grazban tartott nemzetközi konferenciájának a címe: *What is text, really? TEI and beyond (Mi a szöveg valójában? A TEI és azon túl)*. Ez a cím is megmutatja, hogy a digitális filológia, valamint az elektronikus kiadások új textológiaelméleti problémákat hoztak felszínre, amelyek – mint arra Debreczeni Attila is utal fenti cikkében – továbbra is megoldásra várnak. Hans Walter Gablert idézi saját fordításában:

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 következésképpen újramodellezés.¹⁹

Fontos kérdés továbbá az is, hogy a már kidolgozott módszerek mennyire nyugszanak elméletileg jól felépített alapokon.

Mindazonáltal a nemzetközi digitális filológiában több évtized alatt kikristályosodtak bizonyos konszenzusok, és ha még nem is álltak össze egy elméletileg is egységes, viszonylag stabil módszertanná, bizonyos elemeik megkérdőjelezhe-

¹⁸ A szerző egyetlen ilyen hazai közlést ismer, amely azonban az időtállóság követelményének sajnos nem tudott eleget tenni. A Debreceni Egyetemen készült el a Magyar Nyelv Éve alkalmából egy multimédiás összeállítás a Tihanyi alapítólevél és a Tihanyi összeírás nyelvemlékekről (<http://mnytud.arts.klte.hu/kiallitas/intro.html>). Ez tartalmaz digitalizált betűhív szöveget, fordítást, a szöveggel összekapcsolt képet, valamint az egyes szövrányrészletek részletes leírását bibliográfiával, sőt térkép-mellékleteket, valamint hangfájlokat is. A közlés koncepciója közel áll az ideális elektronikus tudományos faksimilékhez (lásd alább), szinte csak a rugalmas keresés lehetőségét hiányolhatjuk. Nagy hátránya viszont a formátuma: a Flash böngészőbővítmény, mely a megjelenítésért felel, napjainkra elavult, nem támogatott, a böngészőkben csak speciális beállításokkal jeleníthető meg, és várhatóan a nem túl távoli jövőben előfordulhat akár olyan felhasználói konfiguráció is, amelyen nem fog futni.

¹⁹ Hans Walter GABLER, „Foreword”, in *Digital Scholarly Editing: Theories and Practices*, eds. Matthew James DRISCOLL and Elena PIERAZZO, XIII–XIV (Cambridge: Open Book Publishers, 2016), XIV. Idézi: DEBRECZENI, „Az elektronikus kiadás színe és visszaja”, 49.

teltenek. Ilyen például az annotáció jelölőnyelvekkel történő megvalósítása. A jelölőnyelvek evolúciója mára egyértelműen kijelölte a leginkább platformfüggetlen és szabványos megoldást, amely a robusztus SGML-től az XML-ig vezetett:

a lassan 20 éve [az idézett cikk megszületésekor, K. Gy.] ISO-szabványként elfogadott SGML (Standard Generalized Markup Language) – szabványos jelölőnyelv dokumentumok belső szerkezetének leírására, beleértve az egyes elemeket jelölő címkék (tagok) definiálásának módját is – segítségével elvben bármilyen dokumentum leírható, függetlenül az azt tároló és megjelenítő számítógépes környezettől, nagyfokú bonyolultsága miatt azonban nem terjedt el sem Európában, sem világszerte a várt mértékben. A technológia viszont a maga nemében nagyszerű és egyedülálló, kár lett volna kiaknázatlanul hagyni. Éppen ezért a webes technológiák fejlesztésével foglalkozó World Wide Web Konzorcium 1998-ban egy új ajánlást – de facto szabványt – jelentett meg, az XML-t

– írta Bíró Szabolcs 2004-ben.²⁰ Az XML tehát az 1986-ban megjelent ISO 8879-es SGML-szabványnak a célszerű szűkítése. Más szabványokkal együtt (például az ISO/IEC 10646-tal, amely a Unicode konzorcium karakterkódolási ajánlásának a szabványa, egy érvényes XML-dokumentum csak ezzel lehet kódolva) képes biztosítani a teljes platformfüggetlenséget. Bíró cikkének megírásakor már készült a legnagyobb nemzetközi szövegekódolási konzorcium, a Text Encoding Initiative újabb ajánlása, a P5 (Proposal 5), amely már XML-alapon tett javaslatokat szövegekódolási eljárásokra. Jelenleg ennek 3.6.0 verziószámú változata van érvényben, amelyet 2019. július 16-án adtak ki. Részben igazat kell adnunk Palkó Gábornak, amikor azt írja:

A digitális filológiában évtizedek alatt egyeduralmukodóvá vált az XML dokumentumleíró szabvány és az arra épülő Text Encoding Initiative (TEI) ajánlás, amelyet kifejezetten bölcsészettudományi használatra, textológus szakemberek közössége hozott létre [...]. A kilencvenes években még komoly viták folytak arról, hogy milyen hátrányai lehetnek a dokumentumok leírására kifejlesztett hierarchikus jelölőnyelveknek, de a TEI végül alternatíva nélkül maradt...²¹

A viták azonban ma is zajlanak. Fennállásának évtizedei alatt a TEI kétségkívül nagy népszerűsége tett szert, valamint a legtöbb textológiai projekt számára kézenfekvőbb megoldás a kitaposott ösvényeken haladni, mint az úttörés

²⁰ Bíró Szabolcs, „A szövegfeldolgozás modern eszközei – az SGML és XML nyelvek”, *Tudományos és Műszaki Tájékoztatás* 51, 10. sz. (2004): 453–459, 454.

²¹ PALKÓ GÁBOR, „Mit jelent a digitális filológia a szemantikus web korában?”, *Magyar Tudomány* 177, 11. sz. (2016): 1316–1322, 1320.

fáradtságos munkáját felvállalni. Ugyanakkor emlékeztetnünk kell Debreczeni Attila felvetésére is:

Abban a tekintetben vita folyik, hogy milyen mértékű lehet és legyen a TEI-XML standardizáció. Vannak vélemények, amelyek olyannyira túlzottan 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.²²

Valójában arról van itt szó, hogy mennyire kell a TEI-nek részletkérdések szabályozásába bonyolódnia. Véleményem szerint is közelebb áll az XML mint jelölőnyelv filozófiájához a rugalmasság, különben megvan a veszélye annak, hogy a TEI az SGML-hez hasonló szörnyeteggé fejlődik, ahol az apróságokat szabályozó részletes elemleírások áttekinthetetlenül és nehézkesen alkalmazhatóvá tennék az ajánlásokat.

A kéziratok tudományos elektronikus kiadásában sem választhatunk tehát mást, mint valamely szabványos jelölőnyelvet, célszerűen a TEI-t. A kiadás elkészítésének a megtervezésekor azonban még egy kérdésre választ kell adnunk: mit fogunk pontosan kiadásnak tekinteni? Maróthy Szilvia fentebb már idézett cikkében az alábbi dilemmákat veti föl:

Mit nevezhetünk tehát kiadásnak az elektronikus közegben? A magában álló TEI XML fájl kiadás-e? A szakirodalom erősen megoszlik a tekintetben, a megjelenítés, felhasználás kritériuma-e a szövegkiadás kiadás voltának. Amennyiben elfogadjuk, hogy a magában álló XML fájl is lehet kiadás, mi húz határt kézirat és publikáció között? Ha pedig az XML fájl megjelenítése(i) által válik publikációvá, hol húzódnak a kiadás határai? E kérdések mind az eredeti problémára vezetnek vissza: a kiadást lezárt egészként képzeljük-e el (mint ami ki van adva). A problémára a digitális filológiai közösségben sem született még konszenzus.²³

A definíciós nehézségeket az okozza, hogy a digitális filológiában határozottan elkülönítik a procedurális és a logikai markupot egymástól. Azaz a megjelenítés csak lehetőség, nem szükségszerűség, vagyis egy XML-fájl publikálása nem a tartalmának és struktúrájának a megjelenítésével egyenlő. Persze könyvtártudományi (vagy ad absurdum tudománytörténeti) szempontból fontos kérdés lehet, hogy mire rakhatunk azonosítószámot, azaz hogy mely informatikai entitás jelentheti az egyértelműen azonosítható digitális objektumot, mondjuk egy kézirat forrás tudományos elektronikus kiadását. Erre talán megadható a válasz:

²² DEBRECZENI, „Az elektronikus kiadás színe és visszája”, 53.

²³ MARÓTHY, „Elektronikus szövegkiadások a könyvtárban”, 299.

a verziószámmal rendelkező szoftver, amelyet fejlesztője bizonyos jól meghatározott kritériumok szerint – mindig ideiglenesen! – lezártnak (jellemzően például új főverzióknak) tekint. Ez a digitális filológiai objektumok esetében akár egy fájl is lehet, de általában egy könyvtár vagy több könyvtár strukturált halmaza. A legjobb példa erre maga a TEI Guidelines, melynek aktuális verziója²⁴ megtekinthető a TEI hivatalos kódrepositóriumában (githubos verziószámmal), és amelynek saját Zenodo DOI-száma van.²⁵ Tehát a Debreczeni Attila által is hangsúlyozott „mediális különmeműség”²⁶ egyik fontos útjelzője lehet, ha az elektronikus kiadásokat a szoftvermérnökök által használt (több közreműködő esetén ún. osztott) verziókövető rendszerek²⁷ segítségével verziószámokkal látjuk el (amelyhez a repositóriumban publikálása pillanatában DOI-szám rendelhető). Debreczeni fel is hívja a figyelmet arra, hogy az elektronikus kiadás nem képernyőn megjelenő könyv:

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. Az így felfogott elektronikus kiadás nem könnyebben használható könyv immár, hanem olyan új kutatási lehetőségek tárháza, amelyek a lineárisan rögzített nyomtatott médiumban elképzelhetetlenek.²⁸

Visszakanyarodva Tószegi fogalmához, az „interpretatív szintű” digitalizáláshoz: a kiadás első lépése az annotálandó szöveg vagy szövegek (szövegváltozatok) előállítás, hiszen az „interpretációnak” (az annotációnak) erre kell épülnie. Ez máris számos elméleti és gyakorlati kérdést felvet. Debreczeni Attila írja: „A szöveg három reprezentációja, azaz a kézírathú átírás vagy kritikai szöveg (*critical text*), az emendált, megtisztított olvasószöveg (*reading text*) és a digitális fakszimile napjainkban már az elektronikus kritikai kiadások meghatározó és megszokott elemei.”²⁹ Tehát az általa „megszokottnak” tekintett digitális fotómásolaton túl két alapvető szövegváltozat felmutatása jellemzi az „elektronikus kritikai kiadásokat”: a kézírathú átírás és az ún. olvasószöveg. A kézírathúság kritériuma régi vitatéma a textológiában, Orlovsky Géza például *A régi magyar*

²⁴ T. E. I. CONSORTIUM, *TEI P5: Guidelines for Electronic Text Encoding and Interchange* Version 3.6.0. (TEI Consortium, 2019).

²⁵ A TEI P5 Guidelines a Zenodón, hozzáférés: 2019.12.19, <https://doi.org/10.5281/zenodo.3413525>.

²⁶ DEBRECZENI, „Az elektronikus kiadás színe és visszája”, 49.

²⁷ A szemantikus verziószámozás egy lehetséges sémája, hozzáférés: 2019.12.19, <https://semver.org/lang/hu/spec/v2.0.0.html>.

²⁸ DEBRECZENI, „Az elektronikus kiadás színe és visszája”, 54.

²⁹ Uo., 52.

textológia helyzetéről című cikkében részletesen foglalkozik a pro és kontra érvekkel.³⁰ Következtetése jó irányelv az elektronikus kiadásokra nézve:

Látható, hogy mindkét szövegkiadási eljárás rendelkezik erényekkel és hiányosságokkal. A kettő közötti választást az anyag természetéből fakadó megfontolásokon túl elméleti meggyőződések is motiválhatják. A döntés sokszor feloldhatatlannak tűnő dilemmáit egy olyan szemlélet haladhatja meg, amely elválasztja egymástól a kritikai kiadás műfajában hagyományosan együtt jelen lévő két funkciót: az archiválást és a szövegprezentációt. Egy ideális szövegkiadás rendelkezésre bocsátaná egyrészt egy mű valamennyi szövegforrásának jó minőségű fakszimiléjét, ezek szigorúan betűhű átiratát, és ezek mellett egy (vagy akár több!) szükség szerint rekonstruált, hangzashű átiratot.³¹

Matthew Driscoll a TEI *Electronic Textual Editing* című kézikönyvében kifejti, hogy a szigorú diplomatikai hűség és a teljesen modernizált átiratok között széles a spektrum,³² a két véglet közötti kevert változatok inkább csak az egyes szövegsajátságok kezelésében mutatnak különbségeket (feloldják-e a rövidítéseket, megtartják-e a nagybetűsítést stb.).

A kézírathűség vagy betűhűség kritériuma a képet is tartalmazó elektronikus kiadások esetében kissé okafogyottnak tűnik (lásd lentebb), illetőleg nincs egyértelmű konszenzus a tekintetben, hogy pontosan milyen transzkripciósi eljárásokkal állíthatók elő a fenti szövegvariánsok (különösen elektronikusan!), a betűhűségnek ugyanis számos szintje létezik. Szentgyörgyi Rudolf a *Tihanyi alapítólevélről* szóló monográfiájának első kötetében foglalkozik a problémával, és megkülönbözteti az ún. betű szerinti és a betűhű átírást (valamint ezektől a paleográfaiilag hűt):

Betűhű – és még nem paleográfaiilag hű – a forrásközlés, ha a forrás szövegét pontosan a forrásban található betűkkel, jelekkel, központosással, egybe- és különírással közöljük. Betű szerinti a közlés, ha a forrásban található betűket a nekik megfelelő mai betűkkel, a forrásnak megfelelő központosással, egybe- és különírással tesszük közzé, a rövidítéseket pedig feloldjuk, de zárójelezéssel vagy más tipográfiai eljárással megkülönböztetjük.³³

³⁰ ORLOVSZKY Géza, „A régi magyar textológia helyzetéről”, *Irodalomtörténet* 85, 3. sz. (2004): 331–344, 333–337.

³¹ Uo., 337. Hozzátehetjük, hogy a „hangzashű átirat” egy újabb átírási mód, az alábbiakban még lesz róla szó.

³² MATTHEW DRISCOLL, „Levels of Transcription”, in *Electronic Textual Editing*, eds. LOU BURNARD, KATHERINE O'BRIEN O'KEEFE and JOHN UNSWORTH, 254–261 (New York: Modern Language Association of America, 2006), <https://tei-c.org/Vault/ETE/Preview/driscoll.html>.

³³ SZENTGYÖRGYI RUDOLF, *A tihanyi apátság alapítólevele: Az alapítólevél szövege, diplomatikai és nyelvi leírása* (Budapest: ELTE Eötvös Kiadó, 2014), 55–56.

Tehát a Debreczeni által olvasószövegnek nevezett változat betű szerinti átírás, ennek elveit a kiadás készítői lefektetik. A nagyobb gond a betűhűséggel (a kritikai szöveggel) van. Ez esetben a paleográfiai hűség és a diplomatikai hűség szintjei között az elektronikus kiadások esetében koncepciónk szerint egy célszerűen megválasztott kritérium alapján dönthetünk: addig megyünk el a betűhűségben, ameddig a szabványos eszközök azt lehetővé teszik, vagyis minden olyan grafematikai sajátosságot visszaadhatunk, amelynek kódolását a Unicode lehetővé teszi. Azért okafogyott valamelyest a Unicode által biztosítottól nagyobb mértékű betűhűséget várni, mert az átiratok egyébként összekapcsolhatók a ma már kellően nagy felbontású digitális képpel, így az eredeti paleográfiai sajátosságai közvetlenül azon tanulmányozhatók. Ennek ellenére vannak olyan projektek, amelyek a Unicode keretein túllépő betűhűséggel dolgoznak: nemzetközi konzorcális keretek között működik például egy olyan, középkori kéziratokkal foglalkozó közösség, a Medieval Unicode Font Initiative (MUFI),³⁴ amely rendszeresen ajánlásokat tesz a Unicode konzorciumnak újabb grafémák kódolására, valamint a Unicode által még nem kódolt grafémáknak a Unicode kódtáblán PUA-nak (Private Use Area) nevezett kódtartományban való kódolására. Ez utóbbihoz a konzorcium saját fontkészleteket is fejleszt. Ezzel azonban sajnos épp a Unicode eredeti célját (az univerzális egységességet és szabványosságot) szüntetik meg. Egyetlen cél indokolhatja a paleográfiailag hű szövegváltozat előállítását: ha kifejezetten ilyen jellegű kutatást szeretnénk támogatni vele, ekkor ugyanis az egyes paleográfiai sajátosságok is kereshetővé vál(hat)nak.

A – Debreczeni kifejezéseivel élve – „emendált és megtisztított” olvasószöveget sokszor normalizáltak is nevezik. A normalizált szövegváltozatok nagyobb mértékben tükröztetnek szerkesztői interpretációt. Az interpretatív mozzanat minden átírás esetében jelen van, és minden esetben szelektálást jelent (Elena Pierazzo megfogalmazásában: „a kiadott szöveg tehát egy az interpretáló-szerkesztő által létrehozott modell, amely a dokumentumban található bizonyos jellegzetességeket kombinál a kutatási célnak megfelelő rendezőelvek alapján”),³⁵ lényeges szempont azonban, hogy ennek eredményeképpen milyen mértékben torzul a szöveg. A különféle normalizálási műveletek elvégzésekor ennek nagyobb a veszélye. Egy ómagyar kori szöveg esetében például gyakran előfordul, hogy egy betű vagy betűkapcsolat olyan hangot jelöl, amely a mai magyarban nincs meg. Ekkor a normalizálás bármilyen formája jelentős torzítást eredményez: ha csak a betűt másoljuk, akkor elsikkad, hogy nem a mai hangértékében szerepel, ha pedig a hangértéket próbáljuk visszaadni, akkor az szükségképpen csak valamely mai betűvel vagy betűkapcsolattal lehetséges egy normalizált szövegben (máskülönben az nem is normalizált), ekkor viszont a betűk mai hangértékével

³⁴ A Medieval Unicode Font Initiative honlapja, hozzáférés: 2019.12.19, <https://folk.uib.no/hnooh/mufi/>.

³⁵ Elena PIERAZZO, *Digital Scholarly Editing: Theories, Models and Methods*, Digital Research in the Arts and Humanities (Farnham: Ashgate, 2015), 44–45. Saját fordítás. K. Gy.

való téves azonosítás torzítja az átírást. Nem teljesen világos, hogy Orlovsky Géza fentebb idézett szövegében mit értett „hangzashű átíraton”, de a nyelvészeti forrásközlések olvasati része, amely valamely fonetikus átírással történik, ezt a problémát hivatott orvosolni – igaz, hogy „népszerű” olvasószövegnek aligha alkalmas. Kérdés, hogy egy tudományos kiadásnak van-e szüksége ez utóbbira. Gyakori még a tudományos szövegkiadások – különösen a nyelvészeti célú közlések – esetében egy szemantikai interpretációt tartalmazó szövegváltozat (értelmezés) előállítás, amely tulajdonképpen mai nyelvre írja át a szöveget (minél régebbi a forrás, annál gyakrabban lehet szükség kihalt szavak, grammatikai jelenségek, maitól eltérő, megváltozott jelentések stb. magyarázatára).

Mindent összevetve valójában a szövegváltozatok materiális és immateriális összetevőinek a kontinuumáról van itt szó:³⁶ az eredeti kéziratot különféle szinteken hűen tükröző átírásoktól (paleográfaiilag hű, betűhű, betű szerinti) a különböző, az interpretációnak nagyobb teret engedő átírásokig (olvasószöveg, normalizált szöveg, olvasat, értelmezés stb.). Mindezek a szöveg különböző dimenzióit ragadják meg (a paleográfiai, ikonológiai, kodikológiaiától a nyelvészeti, szemantikai és irodalmi dimenzióig). A középkori kéziratok tudományos elektronikus kiadásakor célszerűnek látszik a kontinuum minél teljesebb lefedése: fakszimilét is tartalmazó kiadás esetén elegendő lehet egy Unicode-kompatibilis betűhű szöveg az eredeti rövidítésekkel, írásjelekkel stb., egy kidolgozott szabályrendszer alapján emendált betű szerinti szöveg, amelyben a rövidítések feloldva szerepelnek, egy fonetikus átírat, amely akár tükrözhet variációkat is, valamint egy értelmező átírat (ugyancsak célszerűen akár variációkkal). Nagyon fontos szempont, hogy a felhasználó milyen műveleteket végezhet az előállított szövegekkel. A publikált tudományos elektronikus kiadások egyik fontos funkciója például a kereshetőség, amelyet érdemes szem előtt tartani a szövegváltozatok megtervezésekor is. Mivel a speciális Unicode-karakterek gépelése kicsit körülményes lehet, valamint az ortográfiai sokszínűség miatt egy-egy lemma példáinak az előállítása nem szerencsés betűhív vagy betű szerinti szövegben kereséssel (bár az is megoldható). Ezért lehet nagyon hasznos az értelmező átírat, mert abban a szavak mai alakjára kereshetünk.

A következő probléma az előállított szövegváltozatok összekapcsolása. Erre azért van szükség, mert a variációk egyik nagy haszna épp az lehet, hogy a felhasználó által megadott lekérdezések együttesen jeleníthetők meg az egyes átíratokban (illetve akár egy jó minőségű fakszimilén).³⁷ Az együttes lekérdezhetőséghez meg kell határoznunk azt a nyelvi szintet, amelynek a kereshetőségét biztosítani kíván-

³⁶ Uo., 52.

³⁷ Az együttes lekérdezhetőség a fontosabb szempont: nagyon óvatosnak kell lennünk, mert alapelv az elektronikus tudományos kiadásokban, hogy a megjelenítés nem rendelhető alá a strukturális kódolásnak. (Maga az XML épp azért jött létre, mert a weboldalak megjelenítését kódoló HTML jelölőnyelv nem tudta szétválasztani a megjelenítésre és a logikai struktúrára vonatkozó információkat.) A „mediális különeműség”, a hagyományos papíralapú kritikai kiadásoktól való elkülönülés talán egyik leglényegesebb mozzanata ez: a megjelenítő eszköz másodrangúvá válik. Ugyanakkor a majdani felhasználás praktikumája sem kell, hogy elsikkadjon, csak nem válhat a strukturális kódolást vezérlő elvve.

juk. A leggyakoribb megoldás erre a szószintű összekapcsolás, amely alapján a többszavas kifejezésekre történő keresés is biztosítható. Természetesen elő fog fordulni, hogy az egyes szövegváltozatokban a szóalak spacium(ok)kal tagolt kifejezés lesz: ha a scriptor (egységes ortográfia hiányában) nem írt egybe összetett szavakat, ha a toldalékot – vagy éppen toldalékká grammatikalizálódó morfémát – nem írta egybe a szótóvel, ha az értelmezésben csak több szóval lehet kifejezni a forrás szóalakját stb., az a lényeges, hogy ezek a kód szintjén megfelelően legyenek összekapcsolva egymással. Valójában a tudományos elektronikus kiadások nemzetközi gyakorlatában az ilyen szintű összekapcsolásra egyelőre kevés példát találunk.

Az alábbi három egyszerűsített kódrészlet³⁸ a *Halotti beszéd* egy mondatának háromféle szövegváltozatát mutatja be.³⁹

Betűhű:

```
<w xml:id="1">Num</w>
<w xml:id="2">heon</w>
<w xml:id="3">muga nec</w>.
<w xml:id="4">ge</w>
<w xml:id="5">mend</w>
<w xml:id="6">vv</w>
<w xml:id="7">foianec</w>
<w xml:id="8">halalut</w>
<w xml:id="9">evec</w>.
```

Betű szerinti:

```
<w xml:id="1">Num</w>
<w xml:id="2">heon</w>
<w xml:id="3">muga nec</w>.
<w xml:id="4">ge</w>
<w xml:id="5">mend</w>
<w xml:id="6">vv</w>
<w xml:id="7">foianec</w>
<w xml:id="8">halalut</w>
<w xml:id="9">evec</w>.
```

³⁸ A <w> címke a szóalakokat jelzi (word, lásd: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-w.html>). Ezek sorszámot kaptak (xml:id; az itt közöltek nem a valódi kiadás sorszámjai). A sorszám kapcsolja össze a három változat megfelelő szóalakjait. Az írásjelek kódolását a jobb áttekinthetőség kedvéért kivettem.

³⁹ A fonetikus átírás bemutatásától ezúttal az egyszerűség kedvéért eltekintettem. A betűhű és a betű szerinti átírat ez esetben csak a pont nélküli *i* jelölésében tér el egymástól. Vö. Kis Tamás, „A *Halotti beszéd* paleográfiaja és betűhű átíratái”, *Magyar Nyelvtudományok* 56 (2018): 37–56, 40.

Értelmező:

```

<w xml:id="1">Nem</w>
<w xml:id="2">csupán</w>
<w xml:id="3">magának</w>,
<w xml:id="4">hanem</w>
<w xml:id="6">
  <app>
    <choice>
      <orig>ő</orig>
      <reg></reg>
    </choice>
  </app>
</w>
<w xml:id="5">
  <app>
    <rdg>egész</rdg>
    <rdg>minden</rdg>
  </app>
</w>
<w xml:id="7">
  <app>
    <rdg>fajtájának</rdg>
    <rdg>ivadékának</rdg>
  </app>
</w>
<w xml:id="8">halált</w>
<w xml:id="9">evett</w>.

```

Amennyiben szeretnénk biztosítani a 6-os sorszámmal rendelkező *ő* kereshetőségét, akkor meghagyhatjuk a fenti kódolási megoldással, vagy ha a betűhű/betű szerinti átiratunk tartalmaz nyelvészeti annotációt (például a lemmák megadását), akkor az értelmező kódrészletből ki is hagyhatjuk. Ha valaki például az ómagyar E/3. személyes névmásokra akar keresni egy leendő korpuszban, amelyben a *HB*. is szerepelni fog, akkor a névmás mai alakját megadva mindegyik átiratban rátalálhat (vagy az értelmező átiratban szereplő *ő* lesz a találat, vagy a nyelvészeti annotációban megadott lemmája), és akár együttesen (legjobb esetben a képen is) megjelenítheti. A kiadás készítőjének (kódolójának) arra kell ügyelnie, hogy valamilyen módon biztosítsa az eredeti szöveg nyelvi megoldásának a kereshetőségét. Tehát ha nincs nyelvészeti annotáció, azaz nincs megadva a betűhű vagy a betű szerinti átiratban, hogy az *vv* az *ő* névmás alakja, akkor az értelmező részletből sem hagyható el, mert különben nem lenne megtalálható. Látható, hogy variá-

ciók is megadhatók, a fenti példában az értelmező átiratban szerepelnek olyan változatok, amelyek teljesen egyenértékűek egymással: így lehet megadni, hogy az ő egyébként a mai magyarban az ómagyar korival ellentétben nem kötelező (sőt: névelő nélkül agrammatikus, erre utal a kódban a <reg> címke, amely a *regularized* angol szóból ered).

Az elkészítendő kiadás alapvető strukturális kérdése, hogy a szövegváltozatok hogyan jelenjenek meg a kód szintjén. Egy fájlban szerepeljenek? Mindegyik átirat külön fájlba alkosson? Mivel ezek a szövegváltozatok nem a kiadott szöveg genetikai példányai, ezért a TEI ún. párhuzamos szegmentációs eljárása⁴⁰ itt nem megfelelő megoldás. A különböző átiratok egy fájlban kódolására vannak példák a nemzetközi gyakorlatban, de legalább olyan gyakori az is, hogy egy-egy tudományos elektronikus kiadás könyvtárstruktúrába szervezett fájlok együtteseként jelenik meg. Élünk a gyanúperrel, hogy az egy fájlhoz ragaszkodás a hagyományos papíralapú kiadások hatása. Hogy milyen hátrányokkal járhat, azt jól mutatja a MENOTA⁴¹ esete. A Medieval Nordic Text Archive egyike a legszakszerűbb nemzetközi projekteknek, bizonyos megoldásai azonban mégis kérdéseket vetnek fel. Ahogyan kézikönyvük⁴² 4.6. fejezete írja, többszintű átirással készítenek középkori északi germán kéziratokból tudományos elektronikus kiadásokat. Az általuk alkalmazott eljárást Lavrentyev és munkatársai hibrid transzkripciónak nevezik.⁴³ Ennek az a lényege, hogy az átirási rétegeket egy fájlban tartják, a „hibrid” jelző arra utal, hogy egymás mellett szerepelnek a különböző típusú átiratok. Ez azonban általában nem lehetséges szabványosan a TEI keretei között. Az egyik lehetséges megoldása a problémának a TEI saját névtérben alkalmazott kiegészítése, ez viszont szembemegy a konzorcium létrehozásának, működésének egyik legfontosabb céljával: a nemzetközi egységesség követelményével. A MENOTA például háromféle szövegváltozatot kódol, a <me: facs> címkével egy betűhű átiratot (ők „facsimilé”-nek nevezik), amelyben a rövidítések nincsenek feloldva, a <me: dipl> címkével egy betű szerinti (diplomatic, azaz „diplomatikai”), amelyben feloldják a rövidítéseket, továbbá bizonyos egyéb emendálásokat alkalmaznak, valamint a <me: norm>, ami normalizáltat jelent. Így a projekt egyébként igen magas filológiai színvonalát sajnos beárnyékolja, hogy nem tudnak a TEI P5 keretei között maradni.

⁴⁰ A *parallel segmentation method* (párhuzamos szegmentálási eljárás) a TEI Guidelinesban, hozzáférés: 2019.12.19, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/TC.html#TCAPPS>.

⁴¹ A Medieval Nordic Text Archive angol nyelvű honlapja, hozzáférés: 2019.12.19, https://www.menota.org/EN_forside.xhtml.

⁴² Odd Einar HAUGEN, ed., *The Menota handbook: Guidelines for the electronic encoding of Medieval Nordic primary sources*, Version 3.0. (Bergen: Medieval Nordic Text Archive, 2019), <http://www.menota.org/handbook.xml>.

⁴³ Alexei LAVRENTIEV, Yann LEYDIER and Dominique STUTZMANN, „Specifying a TEI-XML Based Format for Aligning Text to Image at Character Level”, 2015, hozzáférés: 2020.02.29, <https://www.balisage.net/Proceedings/vol16/print/Lavrentiev01/BalisageVol16-Lavrentiev01.html>.

Bár szorosabban a megjelenítés kérdésköréhez tartozó téma (amely jelen tanulmányunknak nem tárgya), mégis szükséges szólnunk a képeknek a tudományos elektronikus kiadásokban betöltött szerepéről. A jó minőségű fényképpel való szinoptikus olvasás lehetősége nyilvánvalóan emeli a kiadás színvonalát. A legjobb megoldás viszont az, ha a kép és a szöveg összekapcsolása által megteremtődik annak a lehetősége, hogy a kereséseink és egyéb műveleteink (például a szövegben vagy a képen valamely metaadat eléréséhez kattintás) eredményeit egyszerre jeleníthessük meg a képen és tetszőleges szövegvariánsan (akár valamennyin). Gyakoribb, hogy a forrásdokumentum valamely vizuálisan jól megragadható egysége, jellemzően például a kézirat sora van összekapcsolva az átiratokkal, azonban a szóalapú keresés lehetősége miatt jobb, ha szószinten valósul meg ugyanaz. Annak ellenére, hogy a szabványos technikai feltételei ennek adóttak, még ma sem gyakoriak az ilyesféle kiadások.⁴⁴ A Greta Franzini által 2012 óta szerkesztett *Catalogue of Digital Editions*⁴⁵ keresőjében az alábbi beállításokkal szűrtem meg a több mint 300 kiadást: legyen tudományos (*scholarly*), legyen kiadás, például nem lehet adatbázis (*edition*), legyen TEI-XML, legyenek képek mellékelve, és legyenek a képek összekapcsolva a szöveg(ek)kel (*text-image linking*). A találati lista mindössze 9 tételből állt (ha a szöveg és a kép összekapcsolásának a követelményétől eltekintünk, akkor már 83 tételből áll a lista, vegyesen szerepelnek benne kéziratok és nyomtatványok kiadásai). A találatok között sajnos magyar kiadást nem találunk.

A kliensoldalon (például böngészőben) megjelenítés megvalósításával itt most nem foglalkozom, ki kell térnem viszont arra, hogy a szövegek hogyan kapcsolódnak össze a forrás képével. A képen megjelenő bármilyen vizuális egység kódolható. A TEI P5 módszerének lényege az, hogy a forrás digitális képét zónákra osztja, amelyeknek a koordinátáit összekapcsolja a szöveg megfelelő részeivel, illetőleg azokhoz metaadatokat rendelhet. A zónák lehetnek négyszögletesek, de bármilyen formájú poligon kódolása is megoldható a sarokpontjainak képkoordinátáit rögzítve. A zónák manuális létrehozásához ma már rendelkezésre állnak online is használható eszközök.⁴⁶ Az automatizálás kérdését eddig szándékosan kikerültem (bár vannak ígéretes eredmények, meggyőződésem szerint a megfelelő minőségű digitális filológiai munkához, elsősorban a szövegváltozatok előállításához még sokáig gazdaságosabb módszer lesz szakembert alkalmazni), a nagyobb terjedelmű szövegek képfájljainak a zónákra osztása azonban bizonyos keretek között eredményesen automatizálható. Lavrentyev és munkatársai

⁴⁴ Szinoptikus olvasást lehetővé tevő kéziratkiadás már hazánkban is van: a Digiphil Móricz-levelei például ilyenek, a képek azonban nincsenek összekapcsolva a szöveggel. Hozzáférés: 2019.12.19, <http://digiphil.hu/context/moriczlevelezes>.

⁴⁵ Idézi: MARÓTHY, „Tudományos szövegkiadások a hálózaton”, 617. A katalógus a világháló, hozzáférés: 2019.12.19, <https://dig-ed-cat.acdh.oeaw.ac.at/>.

⁴⁶ Ilyen pl. Marjorie Burghart és Chris Sparks TEI-zonere, hozzáférés: 2019.12.19, <http://teicathuma-num.fr/zoner.php>, ahol tetszőleges fényképet tölthetünk föl, és azt szabványos TEI-annotációval zónákra oszthatjuk.

fentebb már idézett előadásában például egy olyan szoftvert mutatnak be, amely képes automatikus karakterszintű szegmentálásra. A sorokra, illetve szóalakokra bontás pedig egyre több képfeldolgozó szoftver beépített funkciója.⁴⁷ Léteznek olyan projektek is, amelyek képesek a kimeneteiket TEI-XML-ben előállítani.⁴⁸ A szóalakokra szegmentált kép például összekapcsolható a fentebb leírt módon, szószinten összekapcsolt szövegváltozatokkal. Így akár a szóalapú keresés eredménye is megtekinthető egyszerre bármely szövegváltozatban és a forrás képén.

A tanulmányban leírt módszer természetesen csak egyike a kéziratok forrásainak tudományos elektronikus kiadásának. A legfontosabb alapelvei az alábbiakban foglalhatók össze. Az elektronikus kiadás minden eleme szabványos kell legyen. A szövegek szigorúan Unicode-kompatibilis karakterláncok lesznek. Négyféle szövegváltozatot célszerű előállítani: egy betűhű, egy betű szerinti, egy fonetikus és egy értelmező szöveget. A jó minőségű faksimile biztosíthatja a paleográfiai jellegzetességek tanulmányozását – amennyiben a kép össze van kapcsolva az átiratokkal –, így a betűhű változatban megengedhető bizonyos grafematikai egyszerűsítés, de mindent kódolni kell, amit a Unicode lehetővé tesz. Ebben a rövidítések, az összevonások, a javítások, a betoldások, az egybe- és különírás sajátosságai stb. egyaránt tükröződnek. A betű szerinti változat feloldhatja a rövidítéseket, összevonásokat, illetve alkalmazhat egyéb célszerű emendálást. A két, jelentősebb mértékben interpretatív szövegváltozat, az olvasat és az értelmezés a könnyebb olvasást hivatott elősegíteni, valamint – különösen az utóbbi – megkönnyíti a szövegekben keresést. A szövegeket a TEI-XML jelölőnyelvvvel kell annotálni. Az annotáció összekapcsolja a változatokat a szóalakok szintjén, valamint azokat a képpel. Az egyes szövegváltozatok külön fájlban szerepelnek. A TEI-XML-fájlok fejlesztését szigorúan dokumentálni kell valamely osztott verziókövető rendszer (például Git) használatával. A fejlesztés egyes fázisait meg kell tervezni, és meg kell határozni, hogy mit tekintünk alapkiadásnak (1.0 verzióknak). Ehhez DOI-azonosítót (Digital Object Identifier) kell rendelni. Nem alapkövetelmény, viszont kívánatos, hogy a munkálatokhoz nyílt forráskódú és nyílt hozzáférésű szoftvereket használjunk. A tudományos munka átláthatósága, verifikálhatósága érdekében az egész folyamat, valamint a létrehozott produktumok (beleértve a fájlok forráskódját is) ideális esetben nyíltan hozzáférhetők.

⁴⁷ Itt pl. a szerzők összegzik az általuk ismert képfeldolgozó szoftvereket azok funkcióit is felsorolva (Table 1): Nicholas JOURNET, Muriel VISANI, Boris MANSENCAL, Kieu VAN-CUONG and Antoine BILLY, „DocCreator: A New Software for Creating Synthetic Ground-Truthed Document Images”, *Journal of imaging* 3, no. 4. (2017): 62.

⁴⁸ Rory McNICOLL and Timothy MILES-BOARD, „transcriptorium: Computer-Aided, Crowd-Sourced Transcription of Handwritten Text (for Repositories)”, in *10th International Conference on Open Repositories (OR2015)*, 2015; Transkribus. Tanítható HTR (Handwritten Text Recognition, azaz kéziratfelismerő) szoftver. Alkalmos például a kéziratok szöveg sorainak és szóalakjainak az elkülönítésére. Hozzáférés: 2019.12.19, <https://transkribus.eu/Transkribus/>.

FELLEGI ZSÓFIA – PALKÓ GÁBOR

Arany-kéziratok és kritikai kiadások közzététele az Arany János Emlékévb

Manuel Castells, a *hálózati társadalom* nagy hatású teoretikusa a memóriaintézmények szerepét a valós virtualitás összefüggésében vizsgálja. A valós virtualitás az a kulturális szituáció, amikor a kultúra alapvetően és elsődlegesen a digitális médium közvetítésével jut el hozzánk. Ez utóbbi, a digitális médium olyan kommunikációs teret, illetve praxist teremt, amely a kódok széttöredezésével jár, ez pedig – szól az érvelés – megnehezíti, adott esetben ellehetetleníti a sikeres kommunikációt, a közös kódok használatát. Castells szerint ez adja a memóriaintézmények kiemelt szerepét, ezek ugyanis képesek a mindennapi digitális praxisokban széttöredező kódok helyett a megértés lehetőségét biztosító közös tudásteret alkotni.¹

De mi a helyzet akkor, amikor a memóriaintézmények maguk is a digitális közegbe költöznek? A kulturális örökség átalakulása digitális kulturális örökséggé egyrészt a castelli értelemben a kommunikációképesség megőrzésének feltétele a valós virtualitás töredezett világában, másrészt mint transzmedializáció, maga is ki van téve a közeg széttartó effektusainak. Ezen effektusok leírására a digitális kulturális örökség és a digitális humán tudományok összefüggésében Bruno Latour érvelése² kínál vonzó elméleti keretet az időben és térben egymástól távoli, élő és élettelen cselekvő alanyok (aktorok) paradox együttműködéséről.

Amikor a digitális kulturális örökség világában egymással együttműködő aktorok – a közös kulturális mező létrehozása érdekében – szolgáltatásokat építenek, a fenti kettős mediális szituációra kell figyelemmel lenniük. Egyrészt szem előtt kell tartani, hogy csak akkor érheti el egy ilyen „sokszereplős” vállalkozás a célját, ha a valós virtualitás széttartó diskurzusába a szolgáltatás révén minél egységesebb (illetve egységes kódolást lehetővé tevő) produktum kerül be, másrészt folyamatosan kezelni kell az együttműködő aktorok különmeműségét is, mert e különmeműség annál elkerülhetlenebb, minél szélesebb körű együttműködés révén valósul meg a szolgáltatásépítés.

A digitális kulturális örökség és humán tudomány világában az aktorok együttműködése olyan, első hallásra pusztán gyakorlati jelentőséggel bíró kulturális gyakorlatokon keresztül valósul meg, mint a dokumentum- és metaadat-szabványok alkalmazása, a különféle céllal létrejött és különféle szemléleteket implikáló adatleíró nyelvek közötti fordítás vagy megfeleltetés (*mapping*), a külön-

¹ Manuel CASTELLS, „Múzeumok az információs korszakban”, ford. ESZTÁRI Réka, in *Múzeum-elmélet: A képzeletbeli múzeumtól a hálózati múzeumig*, szerk. és előszó PALKÓ Gábor, 30–43 (Budapest: Ráció Kiadó–Petőfi Irodalmi Múzeum, 2012), 30.

² Bruno LATOUR, „A Collective of Humans and Nonhumans: Following Daedalus’s Labyrinth”, in Bruno LATOUR, *Pandora’s hope: Essays on the reality of science studies*, 174–215 (Cambridge, Mass: Harvard University Press, 1999), 183.

nemű adatokat egy platformra hozó aggregáció, illetve az élő aktorok együttműködését kanalizáló technológiai interfész, a kutatókörnyezet. Ezek a praxisok és technikák határozzák meg, a valós virtualításban a kultúra kódjai hogyan vannak (és lesznek) jelen, reflexiójuk azonban nagyon is szegényes a (digitális) bölcsészetben. Ennek egyik oka (Latour egy másik kedvelt kifejezését alkalmazva),³ hogy a gyakran több ezer különműveletű aktor együttműködésével készülő szolgáltatások esetében maguk az adatkezelési (vagy általánosabban: filológiai) gyakorlatok rejtve maradnak: a fekete doboz sötétségében.⁴ Jelen tanulmány megírásának célja, hogy egy konkrét projekt ismertetésén keresztül engedjen bepillantást egy összetett kulturális produktum létrejöttének technológiai folyamataiba, anélkül, hogy e bonyolult (szub)mediális kollaboráció médiaarcheológiai következményeit képes lenne felmérni.

*

Arany János születésének 200. évfordulóján, 2017. március 2-án Nagyszalontán vette kezdetét az Arany János Emlékév, amelynek keretei között számtalan tudományos tanácskozás, kiállítás, kötetkiadás, egyéb, tudományos és népszerűsítő program valósult meg. Az emlékév adott alkalmat annak a nagyszabású digitalizálási projektnek a megkezdésére is, amely magába foglalta az Arany-kéziratok és az Arany János Összes Művei kritikai kiadások online publikálását a nagyközönség számára. Az Arany-kéziratok digitális fakszimiléit összegyűjtő és egységes platformon kereshetővé tevő szolgáltatás (Arany János kézíratainak repozitóriuma), valamint a kritikai kiadások jelölőnyelvi átírata a DigiPhil projekt koordinálásában jött létre.

A DigiPhil (*Tudományos szövegkiadások, bibliográfiák és kutatási adatbázisok online tudástára*), a Petőfi Irodalmi Múzeum és az Bölcsészettudományi Kutatóközpont Irodalomtudományi Intézetének közös vállalkozása minden tevékenység típusában (tudományos szövegkiadások online szolgáltatása, írói bibliográfiák készítése és publikálása, valamint az aggregációs tevékenység) reflektálni igyekszik a bevezetőben említett paradox mediális szituáció támasztotta igényekre (és megfelelő azoknak): a valós virtualitás töredezettségével szemben egységességre, és különműveletű, egymástól időben/térben/személetben távoli humán és technikai aktorok együttműködésének átlátható megszervezésére törekszik.

A kritikai kiadások műhelyeiben a filológiai munkát előkészítendő és segítő, forrásjegyzékek, primer és másodlagos irodalmi bibliográfiák készülnek. Jóllehet az elmúlt évtizedekben ezek szerkesztése már mindenhol számítógéppel történik, a transzmedializáció valójában többnyire még várat magára: a papíron

³ Lásd például: uo.

⁴ A blackboxing kifejezés kapcsán lásd még az alábbi kötet névmutatójában a kifejezés előfordulásait: Pál KELEMEN and Nicolas PETHES, eds., *Philology in the making: Analog/digital cultures of scholarly writing and reading* (Bielefeld: Transcript Verlag, 2019).

készült és a nyomtatásban megjelentetett (irodai szövegszerkesztő programmal készített) bibliográfia egyaránt a nyomtatott médium keretei között marad. Ezzel szemben a könyvtári rendszerek önmagukban, a szemantikus web személtétét is megelőzve, egy adathálózat elemeiként tekintettek rekordjaikra. Helyi névttereket építettek, amelyek publikálásával és egymáshoz kapcsolásával jöttek létre a ma digitális kulturális örökségének alapját képező névtterek, mint amilyen például a VIAF.⁵ A digitális filológia, szemben a nem szöveg típusú digitális objektumokra épülő archívumi gyakorlatokkal, nemcsak a szövegek egészét leíró metaadatokat felelteti meg névtterelemeknek, de a szövegszegmentumok szintjén is végez adatgazdagítást, vagyis hivatkozik a saját személynévtterében és a máshol publikált névtterekben fellelhető besorolási rekordokra. Ilyen jellegű adatgazdagítás nemcsak személy- és helynevek esetében történhet, hanem műfaji fogalmak esetén is érdemes publikált tezauruszokra, ontológiákra hivatkozni. (A DigiPhil például, követve az Europeana példáját, kiadásában a műfaji fogalmakat a Getty Art & Architecture⁶ névtteréből azonosítja.)

Miért van szükség adatgazdagításra? Az egyik nyilvánvaló indok az aggregáció, amely a digitális kulturális örökség talán legradikálisabb és legelőírőbb – de biztosan a legtöbb heterogén aktort integráló – gyakorlata. Olyan nagy hatású intézmények propagálják, mint az Europeana,⁷ amely jelenleg a világ legnagyobb volumenű és hatású aggregátora, közel 60 millió digitális objektummal. Ennek mintájára hozták létre például a Digital Public Library of America-t (DPLA), amely 37 millió digitális objektumot tartalmaz.⁸ Az aggregációs tevékenység során a kisebb helyi gyűjtemények anyagait nemzeti és nemzetközi szolgáltatásokban összeadják és – adatgazdagítás után – kereshetővé és megtekinthetővé teszik.

Az aggregáció azonban nem az egyetlen, talán nem is az elsődleges oka, hogy a digitális filológusok adatmegfeleltetést, szemantikus adatgazdagítást végeznek a digitális objektum, illetve a szövegszegmentumok szintjén is. Ezáltal ugyanis a szabad szavas keresésnél jóval pontosabb és kiterjedtebb kapcsolatrendszer jön létre a szövegek között, amelynek különösen a nem szépirodalmi szövegek (levelezés, naplók) esetén van kiemelt jelentősége. Egy adott kulturális közeg kapcsolati hálózatának mintázatait csak ilyen sűrűségű adatrendszerben érdemes ma vizsgálni. A DigiPhil két projektjében jött létre kulturális mintázatok feltárására alkalmas minőségi adathálózat, a Móricz-naplók és -levelek feldolgozása során, illetve az Arany János kritikai kiadás munkálatainak eredményeképpen. A következőben e munkálatok három területét ismertetjük, előbb a nyomtatott kritikai kiadások jelölőnyelvi kódolásának munkamenetét, majd a kézirat repozitórium

⁵ VIAF: The Virtual International Authority File, hozzáférés: 2020.03.25, <http://viaf.org/>.

⁶ Getty Art & Architecture Thesaurus, hozzáférés: 2020.03.25, <https://www.getty.edu/research/tools/vocabularies/aat/>.

⁷ Europeana Collections, hozzáférés: 2020.03.25, <https://www.europeana.eu/portal/hu>.

⁸ Digital Public Library of America, hozzáférés: 2020.03.25, <https://dp.la/>.

technikai fejlesztését és eredményét, végül a mesterséges intelligencia egy kísérleti alkalmazását a digitális filológia területén.

AZ ARANY JÁNOS KRITIKAI KIADÁS DIGITALIZÁLÁSA⁹

Az Arany János kritikai kiadás digitalizálása a különművelettel együttműködésének egy komplex példája. A korpusz terjedelme (mintegy tizenháromezer könyvoldal) és az Arany- emlékévként időszerűsége nem tette lehetővé, hogy a kódolás munkáját az Irodalomtudományi Intézetben (ITI) működő kutatócsoport Arany-szakértői végezzék, olyan külső munkatársak bevonására volt szükség, akiknek a többsége nem foglalkozott korábban szövegkiadásokkal (esetleg irodalomtudománnyal sem). A munkamenet kialakítása során tehát figyelembe kellett venni, hogy a transzmedializációba bevont munkatársak egy része nem rendelkezett filológiai gyakorlattal, s nem ismerte a TEI XML használatát, illetve a DigiPhil munkamenetét és specifikációját sem. A DigiPhil számára kihívást jelentett azon feltételek megteremtése, amelyek révén úgy állhattak elő jelölőnyelvi átiratok, hogy az Intézet által bevont munkatársaknak¹⁰ nem kellett több hónapos digitális filológiai oktatáson részt venni. Ennek érdekében a digitális bölcsész munkacsoport ergonomikus, egy irodalmi szoftver felhasználói felületét imitáló virtuális kutatókörnyezetet fejlesztett. A grafikus felület mögé „rejtették el” azt az apparátust, amely a voltaképpen jelölőnyelvi átiratot létrehozta.¹¹ Az a törekvés, hogy a kutatók előtt „elrejtjük” a számítógépes kódokat, nem példa nélküli, a TextGrid¹² szolgáltatáson belül is elérhető olyan modul, amely támogatja a TEI XML-ek WYSIWYM („What You See Is What You Mean”) módon történő szerkesztését. Mivel a TEI séma jelenleg közel kétezer oldalt tesz ki,¹³ és számos filológiai jelenség kódolható többféleképpen, használata műhelyenként eltérő, a WYSIWYM szerkesztők pedig igyekeznek a teljes TEI elemkészletet lefedni, a használatukhoz

⁹ A kritikai kiadások listája az alábbi címen érhető el: Magyar Tudományos Akadémia I. Osztályának Textológiai Munkabizottsága, hozzáférés: 2020.03.25, <http://textologia.iti.mta.hu/csv/kiadvany.php?item=16>.

¹⁰ A projekt teljes időtartama alatt az alábbi kollégák dolgoztak a kiadásokon: Csonki Árpád, Horváth-Márjánovics Diána, Hunyadi-Zoltán Emese, Káli Anita, Móré Tünde, Roskó Mira, Sárközi-Lindner Zsófia, Varga Szilvia, Vétek Bence. A DigiPhil csoport (Palkó Gábor, Fellegi Zsófia, Bobák Barbara) munkáját segítették még: Czinkos Zsolt és Metzger Réka (virtuális környezet fejlesztése); Kundráth Péter (megjelenítő fejlesztése); Vincze Tímea és Eőri Lilla (adatgyűjtés, adattisztítás).

¹¹ A DigiPhil korábban, a Móricz Zsigmond-levelezéskiadás kapcsán már próbálkozott olyan leíró környezetet teremteni a Microsoft Office Word szoftverben, amely kiváltja a TEI kódolást, azonban erről a módszerről kiderült, hogy számos előre nem látható hibalehetőséget tartalmaz. A tapasztalatokról a részletes leírást lásd: Cséve Anna, FELLEGI Zsófia és KÓMÁR Éva, „Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadása”, *Digitális Bölcsészlet* 1 (2018): 159–174, <https://doi.org/10.31400/dh-hun.2018.1.227>.

¹² TextGrid: Virtual Research Environment, hozzáférés: 2020.03.25, <https://textgrid.de/>.

¹³ TEI CONSORTIUM, „TEI P5: Guidelines for Electronic Text Encoding and Interchange”, <https://doi.org/10.5281/zenodo.3667251>.

elengedhetetlen a TEI alapos ismerete. A DigiPhil működésének hat éve alatt fejlesztette ki azt a TEI XML specifikációt, amely az összes digitális kiadásban érvényesül, és ami magában foglalja a szövegjellemzők széles körének kódolási mechanizmusát. Ez a specifikáció biztosítja egyrészt a szolgáltatás filológiai és technikai egységét, illetve ez teszi lehetővé, hogy a DigiPhil összes kiadásánál ugyanazt a megjelenítő réteget használja, amely lefedi a különböző kiadástípusok eltérő jellemzőit. Ha a specifikációban rögzítettől eltérő módon használunk egy elemet, akkor az nem, vagy hibásan jelenik meg a DigiPhil oldalán. (Komoly kihívást jelent ugyanakkor, ha egy új kiadás miatt módosítunk a specifikáción, hiszen az hatással lehet a már kész kiadásokra is, illetve ügyelni kell a megjelenítő réteg globális jellegéből adódó hibákra is.)

A DigiPhil a kutatókörnyezet implementálására a digitális filológusok körében az egyik legelterjedtebb szerkesztő eszközt, az Oxygen XML Editor-t használta, amely biztosítja azokat a technikai jellemzőket, amelyek elengedhetetlenek a munkakörnyezet gyors és stabil kialakításához. Az Oxygen eleve rendelkezik alapszintű beépített környezetekkel, úgynevezett keretrendszerekkel (*framework*), a TextGrid-hez hasonlóan a TEI-hez is biztosít egyszerű grafikus szerkesztői felületet, amely azonban nem alkalmas komplex kódok generálására felhasználóbarát módon. A DigiPhil ezért saját keretrendszert hozott létre a szerkesztőn belül, amelyben csak a DigiPhil által meghatározott specifikáció érhető el, és az előre definiált szabályok mentén ellenőrzi, hogy az elkészült fájlok megfelelőek-e. A grafikus réteg mögött lévő XML nem a TEI sémán alapult, hanem egy annál jóval egyszerűbb, magyar nyelvű jelölőket tartalmazó készleten. A munkamenet arra épült, hogy a kódolók a lehető legegyszerűbb felületen dolgozhassanak, a környezetben létrejövő kódot automatizált algoritmusok ellenőrizzék szintaktikai, illetve, ahol ez lehetséges, szemantikai szinten is, majd ezen kódból automatikusan előállítható legyen a publikálásra szánt, a filológiai munka eredményeként értelmezett TEI XML. Erre azért volt szükség, mert egy TEI XML sémának megfelelő dokumentum – különösen annak metaadatokat tartalmazó része – bonyolult szintaxissal rendelkezik, számos olyan kötelező elemet és hierarchikus beágyazást tartalmaz, amelyek használata jelentősen lassította volna a kódolás folyamatát. Olyan köztes, leegyszerűsített kódolási sémát dolgoztak ki, amely automatikusan átfordítható a komplex TEI sémára. A specifikáció kialakításakor a DigiPhil először meghatározta a tizenkilenc kritikai kiadásán belül az egyes kiadástípusokat. Különböző elemkészlettel írható le ugyanis például egy drámafordítás, ahol jelölni kell a színpadi utasítás és a beszélő személyét, vagy a levelezés, ahol a feladót és a címzettet kell leírni. A DigiPhil az egyes kiadástípusokhoz külön környezetet fejlesztett, így a projekt folyamán összesen hét különálló keretrendszer jött létre (drámafordítások, lírai művek, prózai művek, hivatali iratok, levelezés I–II, levelezés III–V, apparátus és jegyzetek). A munka a hivatali iratok és a levelezéskötetek keretrendszereinek kidolgozásával kezdődött, ezeknek a felépítése ugyanis egyszerűbb, mint a szövegváltozatokat is jelölő köteteknek. A kó-

dolók már készségszerűen használták a munkakörnyezetet, mire a bonyolultabb jegyzetapparátussal és több szövegréteggel operáló kötetek leírását megkezdték. A levelezéskötetek esetén azért volt szükség két önálló környezet kialakítására, mert a II. és a III. kötet megjelenése között több, mint húsz év telt el, a kötetek szerkesztési elvei teljesen eltérők voltak, nem lehetett ugyanazzal az specifikációval információvesztés nélkül leírni őket. A kutatókörnyezet ergonómiáját, az irodai szoftverek grafikai felületéhez hasonlóan menüpontok és gombok teremtik meg: egy gomb vagy menü elem – akár a Microsoft Word makró funkciója – olyan műveletsort generál, ami a specifikációnak megfelelő kódrészletet hoz létre. Mindez láthatatlan marad a kódoló munkatárs számára, ugyanis a felhasználói felület grafikus formázott (programkódokat nem tartalmazó) szöveget jelenít meg, a kód jellemzőit grafikai jelekkel, színekkel, elrendezéssel vizualizálva (1. ábra).

A kutatókörnyezet a DigiPhil által előre meghatározott szintaktikai és szemantikai szabályok alapján folyamatosan ellenőrizte a munkatársak által bevitt adatokat, jelölte a kötelezően kitöltendő metaadatmezőket, az egyes jelölők egymásba ágyazhatóságát, a jelölők attribútumainak értékeit, illetve az egyes mezőkbe írt adatok (például a dátumok) formátumát. A leírás során, ha a rendszer kódolási hibát észlelt, jelezte azt a felhasználó számára hibáüzenet vagy figyelmeztetés formájában (2–3. ábra).

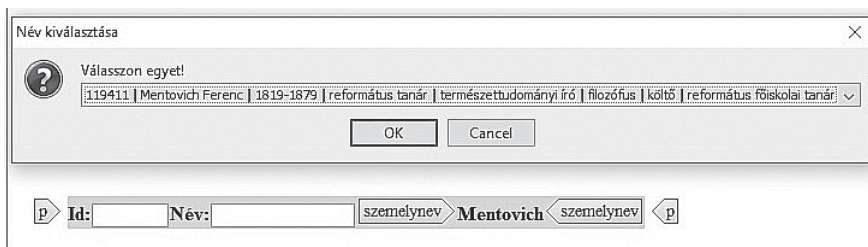
A komplex adathálózatok létrehozásának egyik legfőbb akadálya a munkamenet szempontjából, hogy az adatgazdagítás (nevek, bibliográfiai adatok azonosítása és az azonosítók kódba illesztése) radikálisan lassíthatja a kiadások szövegkódolási folyamatát. Ennek a problémának a megoldására félautomatikus névazonosító funkció került beépítésre a keretrendszerekbe (4. ábra).

A kritikai kiadások névmutatói alapján a DigiPhil névtérben¹⁴ elérhető személynevek azonosítóit és a legfontosabb azonosítást segítő adatait integrálták a rendszerbe. A kódolás során a névazonosító algoritmus személynevek beírásakor azokat automatikusan megfeleltette a névtér elemeinek, a kódoló a rendszer által felajánlott lehetséges egyezések alapján választotta ki a megfelelő besorolási névrekordot, amelynek azonosítója automatikusan íródott be a kódba, ezzel is csökkentve a gépelési hibák esélyét. A kódolás során a bibliográfiai rövidítések jelölésre kerültek, ezek összekötése a Gönczy Monika által a Koha rendszerben épített Arany-bibliográfiával¹⁵ még nem történt meg.

Egy 21. századi digitális filológiai szolgáltatásnak – a digitális médium „illékonyágával” szemben – meg kell felelnie az adatbiztonság legmagasabb elvárásainak. Bár a digitális bölcsészet (nem is olyan távoli) hőskorában Magyarorszá-

¹⁴ A DigiPhil a Magyar Életrajzi Index (Petőfi Irodalmi Múzeum) névtérét építi tovább saját integrált könyvtári rendszerében, amely a Koha szoftvert használja, hozzáférés: 2020.03.25, <http://biblio.digiphil.hu/>.

¹⁵ GÖNCZY MONIKA, „Az Arany János-bibliográfia”, *Értékteremtő tudomány: Az Arany-gyűjtemény bemutatkozása*, Budapest, Bölcsészettudományi Kutatóközpont, Irodalomtudományi Intézet, 2019. november 20.



4. ábra: A félautomatikus névazonosító a környezetben külön ablakban mutatja a találati halmazt, amelyből a kutató választhat

gon is történt komoly adatvesztés,¹⁶ a repozitóriumtechnológia és a verziókövető szolgáltatások általánossá válásával ennek esélye minimalizálható. A transzmedializáció során nagy számú digitális objektum jött létre,¹⁷ amelyek archiválása és változáskövetése volt a munkamenet egyik legkiemeltebb feladata. Mindezt a DigiPhil két szinten kellett, hogy megoldja. Egyrészt el kellett kerülni, hogy a munkamenet ezen része a kódolók heterogén számítógépes környezetétől függjön, másrészt biztosítani kellett a kódolás első (kódolók, munkakörnyezet) és második (a DigiPhil digitális filológusainak feladatai) munkafázisának szétválasztását. A kódolók által kezelt keretrendszer lehetővé tette, hogy a kódokat tartalmazó fájlokat közvetlenül a munkakörnyezetből az erre a célra létrehozott verziókövető szolgáltatásba mentse, így a munka során használt saját számítógépek esetleges biztonságtechnikai vagy egyéb hibái nem veszélyeztették az adatbiztonságot. A verziókövető rendszer követi a szerverre feltöltött kódok módosításait, így a helyi adatvesztés esetén visszaállíthatók a fájlok, a kódoló számítógépén ugyanis csak egy úgynevezett munkapéldány (*working copy*) található. Az eszköz csoportmunka esetén elvégzi a kódokon történt változtatások adminisztrációját. Mivel mindenki saját felhasználói profillal rendelkezik, elkerülhetők a kódolás során egymástól távol „együtműködő” aktorok által módosított kódok okozta adatvesztés. A DigiPhil munkatársai a szerverre feltöltött kódok átalakítását egy attól független rendszerben archiválta.

¹⁶ Az első hálózati betűhű forráskiadás, a Balassa-kódex átírata, faksimilékkel együtt 1993-ban jelent meg online, kísérleti jelleggel, majd 1995-ben CD-változatban is. 1996-ban már tanulmányokkal bővítve jelent meg ismét, azonban 1998-ban megsemmisült. Tóth Tünde, „Online kritikai szövegkiadás Magyarországon az ezredfordulón”, *Helikon* 50, 3. sz. (2004): 417–442. A kiadás további történetét lásd: MARÓTHY Szilvia, „A tudományos elektronikus szövegkiadások feldolgozása”, in *Valóságos könyvtár – könyvtári valóság: Könyvtár- és információtudományi tanulmányok 2018*, szerk. Kiszl Péter és Csík Tibor, 351–356 (Budapest: ELTE BTK Könyvtár- és Információtudományi Intézet, 2018), <https://doi.org/10.21862/vkkv2018>.

¹⁷ Az egyes köteteknél eltérő, hogy egy digitális objektum, azaz egy TEI XML fájl a kötet mekkora egységét fedi le. A levelezésköteteknél például egy fájl egy levelet és a hozzá tartozó jegyzeteket tartalmazza, míg a drámafordításoknál egy fájl egy felvonást tartalmaz, a teljes darabra vonatkozó jegyzetek külön digitális objektumot képeznek. A projekt során összesen több, mint 4200 XML dokumentum jött létre.

A munkamenet függetlenítését a kódolók egyéni számítógépes környezetétől nemcsak a kódok verziókövetése és archiválása érdekében kellett elvégezni, hanem a munkakörnyezet fejlesztése és egységes használata érdekében is. A kódoló munkatársak az Oxygen XML Editor szoftveren keresztül, egy szerverre mutató link segítségével telepítették a keretrendszert, ami ezt követően minden újraindításkor ellenőrizte az elérhető csomag verziószámát, és amennyiben változtatás történt, automatikusan frissítette a rendszert. Ez lehetővé tette, hogy minden munkatárs – függetlenül attól, hogy milyen számítógép(ek)en dolgozott – mindig ugyanazt a kutatókörnyezetet használja, költséges és időigényes hardveres támogatás igénybevétele nélkül.

A projektbe bevont munkatársak a kutatókörnyezet segítségével közel másfél év alatt dolgozták fel az Arany kritikai kiadások köteteit. Ezt követően a DigiPhil munkatársai a létrejött fájlokból egyedi fejlesztésű stíluslapok segítségével előállították a publikálásra szánt TEI XML fájlokat, és feltöltötték lektorálásra a szolgáltatást biztosító repozitóriumba. A munkatársak az „éles” szolgáltatással mindenben megegyező tesztkörnyezetben ellenőrizték a kiadások szövegét, míg a digitális filológusok a TEI XML fájlokat vizsgálták át, a javításokra már a TEI XML fájlokban került sor.

A kötetek publikálásához szükség volt a specifikáció implementálására: a megjelenítő réteg fejlesztésére, hiszen a kritikai kiadássorozat kötetei több, mint fél évszázad (hatvanöt év) filológiai munkáját – és az azalatt történt szemléletváltozást – foglalják magukba. A szolgáltatás ünnepélyes bemutatására 2019 novemberében került sor az Irodalomtudományi Intézetben.¹⁸

Bár a digitalizálási munka lezárult, a virtuális kutatókörnyezetek építése és alkalmazása folytatódik. A DigiPhil szakértői a projekt során szerzett tapasztalatokra építve olyan környezetet fejlesztenek, amelynek az alapját már nem egy leegyszerűsített köztes XML séma adja, hanem a TEI specifikációnak mindenben megfelelő, az adott kiadásra alkalmazott végleges séma. Ezzel lehetőség nyílik arra, hogy a digitális kritikai kiadások a munkafolyamat kezdőpontjától egységes és szabványos formátumban készüljenek, és publikálás után is folyamatosan javíthatók és bővíthetők legyenek a létrehozás folyamatában felhasznált keretrendszerrel azonos környezetben. A folyamatban lévő Arany kritikai kiadások már ebben a környezetben készülnek, az S. Varga Pál által szerkesztett *Arany János Kisebbségi Költemények 3. (1860–1882)* jelölőnyelvi átirata, a nyomtatott kiadás javított, digitális változata is az új környezetben születik meg.

¹⁸ *Értéktéremtő tudomány: Az Arany-gyűjtemény bemutatkozása*, Budapest, Bölcsészettudományi Kutatóközpont, Irodalomtudományi Intézet, 2019. november 20.

ARANY JÁNOS KÉZIRAT-REPOZITÓRIUM

Az Arany kritikai kiadás során Arany-filológusok, digitális bölcsészek és kódoló munkatársak működtek együtt egy olyan kutatókörnyezet biztosította digitális térben, amely maga is különféle forrású, korú és elhelyezkedésű, heterogén technikai aktorok (kliens és szerver számítógépek, szoftverkomponensek, szabványok, sémák, kódnyelvek) összjátékaként írható le. A következőkben egy térben és időben még heterogénebb intézményközi kollaboráció felépítését és munkamenetét ismertetjük, amely talán a korábbiaknál is pontosabban példázza a digitális kulturális örökség széttartó forrásokra épülő, de egységes felületeket képező kortárs gyakorlatait.

Az Arany-kéziratok online publikálását célzó projekt kezdeményezője és koordinálója az Irodalomtudományi Intézet (ITI) volt, a résztvevő intézmények pedig a nagyszalontai Arany János Emlékmúzeum, a nagykőrösi Arany János Kulturális Központ, a Magyar Tudományos Akadémia Könyvtár és Információs Központ (MTA KIK), a Petőfi Irodalmi Múzeum (PIM) és az Országos Széchényi Könyvtár (OSZK) voltak.

Az Arany-kéziratkereső létrehozásának alap gondolata az volt, hogy a fizikailag szétszórt, különböző intézményekben fellelhető Arany János kéziratok egyetlen online térben váljanak elérhetővé a kutatók és a szélesebb közönség számára egyaránt, a metaadatok szerint kereshető formában. A kezdeményezésnek különös hangsúlyt ad a tény, hogy a magyar irodalom kéziratok örökségén (számos kisebb gyűjtemény mellett) három nagy intézmény osztozik (OSZK, MTA-KIK, PIM), de sem a tömeges digitalizálás, sem a kéziratok le- vagy átírása, sem pedig az online közzététel szintjén nem működtek eddig együtt.

Az Arany kézirat-repozitórium létrehozásához meg kellett teremteni azt az intézményi együttműködést, amelynek során a kéziratokat gondozó intézmények közös elvek mentén készítik el a kéziratok digitális másolatait és összehangolt módszerekkel végzik el a kéziratok leírását. A kéziratkereső specifikációját a DigiPhil-csoport alakította ki az együttműködő intézmények könyvtárinformatikus szakértőinek segítségével.

A digitalizálás első lépéseként az Országos Széchényi Könyvtár és az Irodalomtudományi Intézet munkatársai elvégezték az Arany János Emlékmúzeumban és az Arany János Kulturális Központban található kéziratok digitalizációját. A folyamat során nagy felbontású képfájlok jöttek létre, majd ezekből a fájlokból készültek el a publikálásra szánt nézőképek. Az OSZK, a PIM és az MTA KIK a saját gyűjteményeikben fellelhető kéziratok digitalizálását a belső szabályzataiknak megfelelően végezték el, a képfájlokat a saját rendszereikben tárolták.

A közös szolgáltatás kialakításának egyik sarkalatos pontja az aggregált tartalmak egységesítése, illetve a képfájlok elhelyezése volt. Követendő példának az Europeana aggregációs modellje kínálkozott: az Europeana csak a metaadatok összegyűjtését végzi el, a memóriaintézmények gyűjteményi elemeinek digitális szurrogátumait maguk az intézmények teszik közzé: az aggregált metaadatok

egyike a digitális objektumra mutató link. Az Europeana a beérkező metaadatok egységesítésével sem foglalkozik, azokat saját formátumában fogadja csak el (Europeana Data Model, EDM), az adatszolgáltató intézmények saját felelőssége a metaadatok konverziója EDM formátumra.

A fenti modell sem a digitális objektumok, sem pedig a metaadatok kezelése szintjén nem volt alkalmazható a projekt számára. A határon túli gyűjtemények nem rendelkeztek repozitóriumi kapacitással, vagyis a digitális objektumok szolgáltatását is meg kellett oldani. A magyarországi közgyűjtemények ugyanakkor eltérő megoldásokat alkalmaztak a digitális faksimilék közzétételére, amelytől nem akartak eltérni, emellett különböző intézményi szabályzatok alapján kezelték a digitális objektumaikat. A probléma heterogén technikai megoldásokat követelt meg mind a metaadatok, mind a faksimilék kezelésében.

A digitális objektumok szolgáltatásához a repozitóriummal nem rendelkező határon túli partnerek számára a Bölcsészettudományi Kutatóközpont virtuális szerverén a DigiPhil egy olyan repozitórium szoftvert telepített, amely képes a digitális faksimilék tárolására. Ehhez a repozitóriumhoz nem készített nyilvános felhasználói felületet, az abban található objektumok közzététele egyedül a közös keresőfelületen történik. Ugyanezt a technikai megoldást (és ugyanezt a DSpace nevű szoftvert) alkalmazta az OSZK is saját szerverén. A faksimilét az MTA KIK az EPrints szoftverre épülő saját REAL rendszerében helyezte el, a partnerek közül egyedül kézirategységként egyetlen PDF-be egyesítve, a PIM pedig a HunTéka részeként működő JaDoX dokumentumkezelő szoftverben helyezte el a képeket. Az MTA KIK és a PIM kéziratára az átadott rekordokban tüntette fel a képek elérési útvonalát.

Mivel a projekt elsődleges feladata a kéziratokhoz készülő közös metaadat-kereső-felület kialakítása volt, figyelembe kellett venni azt a tényét, hogy bár a résztvevő intézmények a könyvtári rendszereikben szabványos formátumban írják le a rekordokat, az egyes könyvtári rendszerekben használt metaadatsémák eltérők lehetnek. A keresőfelület kialakításakor a résztvevők a MARC21 könyvtári adatsere formátum használata mellett döntöttek. Mivel a két határon túli intézmény nem rendelkezett megfelelő technikai háttérrel, kézírataikat a DigiPhil könyvtári rendszerében írták le az OSZK (Fekete Eszter és Rózsafalvi Zsuzsanna) és az ITI (Antal Alexandra és Bátori Anna) munkatársai. A Koha által használt metaadatformátum ugyan MARC21, azonban a DigiPhil korábban még nem írt le kéziratokat, ezért új metaadatspecifikáció kialakítására volt szükség. Az OSZK, annak érdekében, hogy a kéziratok leírása egységes legyen, nem a saját, HUNMARC alapú rendszerükben, az Amicusban, hanem egy új, erre a célra telepített Koha példányban dolgozta fel a kézírataikat. A DigiPhil metaadatspecifikációját a PIM könyvtárinformatikus szakértője, Mohay Anikó készítette elő, majd ez került kiegészítésre az OSZK munkatársa, Fejes Judit javaslataival. A DigiPhil és az OSZK Koha rendszereiben a kéziratok leírása ugyanazon specifikáció alapján zajlott. Mivel a feldolgozást végző munkatársak közül nem minden-

ki rendelkezett tapasztalattal a könyvtári rendszerek használatában, az adatfelvitel megkönnyítésének érdekében Mohay Anikó magyar nyelvű űrlapokat hozott létre, amelyek nem mutatták a teljes MARC21 adatmezőkészletet, csak a leíráshoz szükségeseket, emellett sok mező esetében (például proveniencia) pedig előre meghatározott listából választhattak a kollégák. A két Koha rendszerben leírt rekordok esetében, ahol lehetséges volt, a leíró kollégák adatgazdagítást végeztek. Egy előzetesen, a kritikai kiadások helynevei alapján készített listából felvitték a helynevekhez a GeoNames névterében található azonosítót, tárgyszavakkal látták el a rekordokat, valamint azokat a kéziratokat, amelyek valamely kritikai kiadásban megjelentek, összekötötték a DigiPhil oldalán publikált digitális kiadással. Az MTA KIK már 2016 folyamán átállt a MARC21 adatformátumra,¹⁹ a PIM saját HunTéka rendszeréből szintén ebben a formátumban nyerte ki és adta át a publikálásra szánt rekordokat. Az MTA KIK a saját gyűjteményéhez tartozó anyagok metaadatolását az Aleph könyvtári rendszerben végezte.

A közös adatsere formátum használatának ellenére felmerültek olyan jelölési eltérések, amelyeket egységesíteni kellett a rekordok közös rendszerbe való betöltése előtt. Az MTA KIK által szolgáltatott rekordokban a személyneveket leíró 100-as (szerző), 600-as (róla szóló) és 700-as (egyéb személynév) mezőkben a vezetéknevet és keresztnévet vesszővel választották el, míg a többi intézmény által szolgáltatott rekordokban csak a nem magyar személynevek esetén alkalmaztak hátravetést. Annak érdekében, hogy a keresőfelületen ezek az értékek egységesen szerepeljenek, az MTA KIK-től kapott rekordokban félautomatikus módon, reguláris kifejezések segítségével a magyar személynevekből el kellett távolítani a szeparátorokat. Ugyanezen mezők \$d almezőjében, amelyben a személy születési és halálozási dátuma adható meg, a többi intézménnyel ellentétben a DigiPhil leírásában az évszám zárójelbe került, ezeket szintén egységesítették. Az egységesítésnél kiemelt szempont volt, hogy minimális változtatások történjenek a rekordokon és azokat a lehető legrészletesebben dokumentálják, hogy a műveletek bármikor megismételhetők legyenek, hiszen ezek a változtatások nem kerültek vissza az eredeti rendszerekbe, egy esetleges frissítés során újra végre kell hajtani őket. A mezők kitöltésének eltérésein túl előfordultak olyan esetek is, amikor a gyűjteményekben ugyanazt az információt más adatmezőbe vitték fel. Ilyen például az egyedi azonosító, ami a DigiPhil és az OSZK által leírt rekordok esetében a 999-es mezőbe került, míg az MTA KIK a 001-es mezőt használta ugyanerre a célra. A különbség a használt könyvtári rendszerekből fakad, a Koha sajátosság, szemben az Aleph-el, hogy a kiosztott azonosítót nem a 001-es mezőbe helyezi. Hasonló a helyzet a tárgyszavakkal: míg a DigiPhil és az OSZK a leírás során a tárgyszavakat a 650-es mezőbe vitte fel, addig az MTA KIK ugyanerre a 653-as mezőt használta. Az ilyen jellegű eltéréseket a kereső algoritmus finomí-

¹⁹ GYURICZA Andrea, „Szabványos metaadatok jelentősége a kooperatív szolgáltatásokban: Átállítás HUNMARC-ról a MARC21 szabványra a WorldCat-csatlakozás előkészítéseként”, *Könyvtári Figyelő* 63, 2. sz. (2017): 218–225.

DIGIPHIL PROJEKTEK SZERZŐK MŰVEK

Kézirat kereső

Keresés...

Rendezés: Dátum (növekvő) ▼

Szerző/Közreműködő

Cím/Megnevezés

Dátum

Dedikáció

Hely

Megjegyzés

Műfaj

Róla szóló

Proveniencia/Származás

Tárgyszó

ID

5. ábra: Keresőmezők az Arany-kéziratok repozitóriumában

tásával oldották meg, például a tárgyszavak esetében a rendszer egyszerre figyelni mindkét mezőt.

Az Arany-kéziratok repozitórium keresőfelülete egyszerre támogatja a szabad szavas és facettás keresést a metaadatok között. A „Részletes keresés” menüpont alatt található azok a keresőmezők, amelyek az egyes metaadatmezőket reprezentálják, ezekben szabad szavasan és legördülő listából való választással egyaránt kereshet a felhasználó (5. ábra).

A facettás kereső használatakor a felhasználó egyszerre több metaadatmezőben végezhet szűkítést, továbbá az egyes listák esetében is egyszerre több értéket jelölhet ki, ekkor a szűkítés során a keresőmotor minden kiválasztott értéket figyelembe véve jeleníti meg a találatokat (6. ábra).

A keresőrendszer kialakításakor elsődleges szempont volt egy olyan, nem robosztus, nyílt forráskódú eszköz választása, amelynek hosszú távú fenntartása, fejlesztése biztosított, emellett az adatbetöltés, frissítés felhasználóbarát módon történhet. A választás az eXistDB nevű adatbázis-kezelő szoftverre esett, amely képes nagy mennyiségű XML formátumú fájl tárolására és kezelésére, miközben a felhasználói felület egyszerűen testre szabható. A DigiPhil projekt ugyanezt az

Szerző/Közreműködő	Arany János 1817-1882
Cím/Megnevezés	Bármely
Dátum	ÉÉÉÉ - ÉÉÉÉ
Dedikáció	Bármely
Hely	Bármely
Megjegyzés	
Műfaj	vers
Róla szóló	Bármely
Proveniencia/Származás	Arany János Emlékmúzeum, Nagyszaló
Tárgyszó	<input type="text" value="Search"/>
ID	<input checked="" type="checkbox"/> Arany János Emlékmúzeum, Nagyszalonta <input type="checkbox"/> Magyar Tudományos Akadémia Könyvtár és Információs Központ <input checked="" type="checkbox"/> Nagykőrösi Arany János Kulturális Központ <input type="checkbox"/> Országos Széchényi Könyvtár <input type="checkbox"/> Petőfi Irodalmi Múzeum

Összesen: 8 találat

6. ábra: Szűrés az Arany János Nagykőrösön és Nagyszalontán található verskézírataira

eszközt használta a *Móricz-levelek* keresőjének fejlesztéséhez is.²⁰ A szolgáltatás kialakításakor fontos szempont volt, hogy a magyarországi közgyűjtemények maguk szolgáltatassák a digitális faksimiléiket, a keresőbe csak a metaadatok kerüljenek feltöltésre. A felhasználói felület tervezésekor sikerült megoldani, hogy a DSpace rendszerekben tárolt fájlok magán a közös keresőfelületen kerüljenek megjelenítésre.

A programban résztvevő intézmények a projekt végére összesen 2003 rekordot szolgáltatottak. Bár a projekt véget ért, az MTA KIK folyamatosan javítja és egységesíti a katalóguscédulákból létrejött rekordjait, valamint feldolgozza a még digitalizálatlan kéziratokat, így ezekkel tovább bővül majd a szolgáltatás. A projekt során kialakított munkafolyamat alkalmas arra, hogy a jövőben hasonló módon dolgozzuk fel és publikáljuk más szerzők kéziratok hagyatékát. A kialakított heterogén aggregációs modell azt is lehetővé teszi, hogy a szolgáltatáshoz különféle technológiai és intézményi feltételekkel csatlakozhassanak memóriaintézmények és kutatócsoportok.

²⁰ A *Móricz-levelek* keresőfelülete, hozzáférés: 2020.03.25, <http://digiphil.hu/mzs-lev-kereses/index.html>.

KÉZIRATÁTÍRÁS GÉPI INTELLIGENCIA RÉSZVÉTELÉVEL

Az eddigiekben olyan szolgáltatásokat ismertettünk, amelyek különmemű aktorok együttműködésében jöttek létre, és amelyekben a humán közreműködők kollaborációját valamely magában is heterogén virtuális kutatókörnyezet vagy technológiai platform kanalizálta. Az ismertetett szolgáltatásokban tevékeny (emberi és nem-emberi) aktorok köre ugyanakkor a megszokott vagy szabályozott keretek között maradt mind a humán, mind a technikai oldalon: digitális bölcsészek, szabványok, széles körben használt szoftvereszközök működtek közre. Ugyanakkor a digitális kulturális örökség, illetve filológia területén számos olyan gyakorlat létezik ma, amely a fentieknél radikálisabban nyitja meg az aktorok körét. Az egyik ilyen, mind szélesebb körben alkalmazott praxis a kéziratok átírásának közösségi (*crowdsourcing*) kiterjesztése. A már klasszikusnak számító Bentham Project²¹ mellett érdemes az Europeana hasonló, még szélesebb kört megszólító kezdeményezését is szóba hozni.²² Az ismeretlen, vagyis kontrollálatlan kompetenciával (és intenciókkal) rendelkező szereplők bevonása szigorúan filológiai projektekbe különféle technológiai aktorok közreműködésével kontrollálható (például automatikus kolláció révén), de a magukat humán közreműködőnek álcázó rosszindulatú kódok kiszűrését is gépi aktorok végzik.²³ Még radikálisabb azonban az a gyakorlat, amelyik a gépi intelligenciát a korábban kizárólag emberi szereplők kompetenciakörébe utalt feladatokba integrálja. Ilyen terület a mesterséges intelligencia felhasználása a kéziratok átíratainak elkészítésében.

Az optikai karakterfelismerés (OCR) gyakorlatait már jó ideje alkalmazzák a kulturális örökség digitalizálásában, legtöbbször olyan hibrid, a digitális és az analóg médium határán egyensúlyozó formátumok létrehozásában, mint amilyen a kétrétegű dokumentum, amelyik a nyomtatott médium elrendezését (*layout*) a kereshetőség igényével párosítja. Köztudott, hogy az OCR technológia kevésbé alkalmas kézírásos dokumentumok átírására, az azonban kevésbé ismert tény, hogy már az '50-es évek közepétől, párhuzamosan a nyomtatott szövegek gépi átírásának fejlesztésével, a kéziratok anyagok gépi átírására is történtek kísérletek.²⁴ A kézírás-felismerés (*handwritten text recognition*, HTR) alkalmazása a pénzügyi és banki ágazatokra korlátozódott (például levélcímzés felismerése, csekkek elemzése). A tömeges gyűjteményi digitalizálás hozta magával azt az igényt, hogy a kéziratok kikerülve, vagy legalábbis előkészítve a humán aktorok

²¹ A Bentham Project oldala, hozzáférés: 2020.03.25, <https://www.ucl.ac.uk/bentham-project/>.

²² Enrich Europeana, hozzáférés: 2020.03.25, <http://transcribathon.eu/>.

²³ Jussi PARIKKA and Tony D. SAMPSON, „Censored”, in *The spam book: on viruses, porn, and other anomalies from the dark side of digital culture*, eds. Jussi PARIKKA and Tony D. SAMPSON, 213–215, Hampton Press communication series: Communication alternatives (Cresskill, N. J.: Hampton Press, 2009).

²⁴ Guenter MUEHLBERGER, Louise SEAWARD, Melissa TERRAS, Sofia ARES OLIVEIRA, Vicente BOSCH, Konstantinos ZAGORIS et al., „Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study”, *Journal of Documentation* 75, no. 5. (2019): 954–976, 966, <https://doi.org/10.1108/JD-07-2018-0114>.

végezte átírás lassú és fáradságos folyamatát, ugyanúgy kereshetővé és gépek számára is feldolgozhatóvá válnak, mint a nyomtatott és OCR-ezett szövegek.

A HTR technológia legnagyobb kihívása a kézírás egyedisége, erre jelent megoldást a gépi tanulás friss trendje, a neuronháló (Artificial Neural Network, ANN) technológia alkalmazása a HTR folyamatában. Amennyiben egy szerzőtől megfelelő mennyiségű átírat és faksimile áll rendelkezésre, a gép képes lehet megtanulni a „kézre” jellemző stílusjegyeket és létrehozni egy olyan modellt, amely felismeri, kiolvassa a szerző kézírását. Minél több átírat és nagyobb számítási kapacitás áll rendelkezésre, a modell annál pontosabb lesz.

2016-ban hozta létre az Innsbrucki Egyetem *Digitalisierung & Elektronische Archivierung*²⁵ kutatócsoportja a READ (*Recognition and Enrichment of Archival Documents*)²⁶ projekt keretében a Transkribus szolgáltatást.²⁷ A Transkribus volt az első olyan platform, amely nyílt hozzáférést biztosított egy felhasználóbarát felülettel rendelkező és neuronhálón alapuló HTR eszközhöz.

Az ELTE Digitális Bölcsészeti Központ az ITI-vel együttműködve kezdett kísérletezni Arany János hivatali iratainak felismertetésével a Transkribus szolgáltatásban. A Transkribus rendelkezett ugyan kész kézírás-felismerő modellekkel, azonban azokat angol és német nyelvű kéziratokon tanították be, így magas hibaszázalékkal működtek az Arany-kéziratokon. A Transkribus lehetőséget biztosít arra, hogy saját modellt hozzunk létre, így a Központ munkatársai száz kéziratlap átíratát készítették el, amelyből a Transkribus algoritmusai létrehozták az első Arany-kézírás modellt. Az első modell ellenőrzése során kiderült, hogy karakterek szintjén mintegy 30%-os hibaszázalékkal (Character Error Rate, CER) működött. A második modell előkészítésében és tesztelésében már a DigiPhil munkatársa, Bobák Barbara és az ITI munkatársa, Gábori Kovács József vettek részt.²⁸ Újabb száz, Arany által írt kéziratlap került átírásra, így az új modell már kétszáz kéziratlap átírata alapján készült el. A Transkribus a második modell elkészülte előtt javított a modellt készítő algoritmuson, a Google által fejlesztett Tensorflow alkalmazásával.²⁹ Ennek következtében jelentősen javult a neuronháló teljesítménye. Ahogyan Bobák és Gábori írják,³⁰ a második modell hatékonysága erős javulást mutatott, a CER a felére csökkent. A Transkribus saját tapasztalatai szerint a modell tanításával

²⁵ Digitalisierung & Elektronische Archivierung (DEA), hozzáférés: 2020.03.25, <https://www.uibk.ac.at/germanistik/einrichtungen/dea.html>.

²⁶ Recognition and Enrichment of Archival Documents, hozzáférés: 2020.03.25, <https://read.transkribus.eu/about/>.

²⁷ Transkribus, hozzáférés: 2020.03.25, <https://transkribus.eu/Transkribus/>.

²⁸ BOBÁK Barbara és GÁBORI KOVÁCS József, „Kézírásfelismerés Arany János levelein”, in *Workshop 2019*, szerk. TÍCK József, KOKAS Károly és HOLL András, 38–44 (Budapest: HUNGARNET Egyesület, 2019), <https://doi.org/10.31915/NWS.2019.5>.

²⁹ MUEHLBERGER, SEAWARD, TERRAS, ARES OLIVEIRA, BOSCH, ZAGORIS et al., „Transforming Scholarship...”, 966.

³⁰ BOBÁK és GÁBORI KOVÁCS, „Kézírásfelismerés Arany János levelein”, 6.

a CER akár 5% alá csökkenthető.³¹ Mivel a modellt csak az Arany által írt kéziratokon tanították be, jelenleg nem alkalmas arra, hogy más szerző kézírataiból jó minőségű átiratokat készítsen, azonban a modell tanítóanyagának bővítésével lehetővé válhat általánosabb érvényű modellek megalkotása is.

A Transkribus szolgáltatás egyik nagy előnye, hogy az átírás eredményét, legyen az gépi közreműködés eredménye vagy humán aktor által javított, a hétköznapi formátumok mellett (TXT, PDF) olyan TEI XML formátumban is exportálhatjuk, amely tartalmazza az átirat és a faksimile kép kapcsolatát. Ezáltal lehetőség nyílik arra, hogy a gépi intelligencia közreműködésével meglepően rövid idő alatt olyan digitális forráskiadásokat készítsünk, amelyek – emberi közreműködéssel – kritikai kiadások alapjául szolgálhatnak, hiszen a feldolgozás az első lépésektől szabványos környezetben történik.

³¹ MUEHLBERGER, SEAWARD, TERRAS, ARES OLIVEIRA, BOSCH, ZAGORIS et al., „Transforming Scholarship...”, 962.

SZEMLE

SZÉNÁSI ZOLTÁN

Digitális átalakítás

*A tudományos szövegkiadás lehetőségei a weben**

Ha a „digitális filológia” fogalmát a jelzett szó, a „filológia” felől közelítjük meg, akkor az irodalom speciális eszközökkel és saját metodikával dolgozó részterületeként foghatjuk fel, a tudományos munkát segítő eszköz alkalmazásánál azonban mindenképpen többről van szó. A „digitális” jelző egyértelműen kijelöli a terület interdiszciplináris jellegét, melynek értelmében a digitális filológiát az ún. digitális bölcsészet (*digital humanities*) egyik ágaként határozhatjuk meg, amennyiben a digitális bölcsészetet a számítástechnika és a kulturális örökség metszéspontjában, a művészeti alkotások és a rájuk irányuló bölcsészettudományi kutatások digitális átalakításán dolgozó tudományos professzióként értjük.¹ Andrew Oliver az informatika textológiai alkalmazásának lehetőségeit mérlegelő tanulmányában a nyolcvanas évek végén úgy tekintett a számítógépre mint a textológiai tevékenység új és hatékony eszközére, mely – mint írja –

igazi, intellektuális szerepébe helyezte vissza a textológust. Azzal, hogy megkönnyíti a feladatát és szigorú önfegyelemre szorítja – hiszen a számítógép nem tűr pontatlanságot –, a gondolkodás szabadságával ajándékozza meg, amit azelőtt a gyakorlati teendők nyomása alatt ritkán engedhetett meg magának.²

Hasonló optimizmussal vizsgálja a digitális médiumok használatának a filológia diszciplináris pozíciójára gyakorolt hatását az ezredforduló után Jerome McGann is, és a szövegtudomány irodalom- és kultúratudományt megalapozó szerepének visszatérését prognosztizálja.³ Kérdés azonban, hogy az új eszközök használatbavételén, s a használat következményeként a régi textológusi erények – miként például az Oliver által emlegetett pontosság – megerősítésén túl milyen

* A tanulmány a Bolyai János Kutatási Ösztöndíj támogatásával készült.

¹ Vö. Julianne NYHAN and Andrew FLINN, „Introduction”, in Julianne NYHAN and Andrew FLINN, *Computation and the Humanities: Towards an Oral History of Digital Humanities*, 1–19 (Cham: Springer International Publishing, 2016), 1, <http://dx.doi.org/10.1007/978-3-319-20170-2>.

² Andrew OLIVER, „Mikroinformatika és textológia”, ford. FARKAS Ildikó, *Helikon* 35, 3–4. sz. (1989): 412–420, 420.

³ Jerome J. MCGANN, *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction* (Cambridge, MA–London: Harvard University Press, 2014), 20.

más eredményei vannak a számítógép alkalmazásának a filológia területén. A tudományos igényű hálózati szövegkiadás magyarországi eredményeinek áttekintésekor az elsődleges kérdés számomra az, hogy milyen lehetőségeket kínál a webes médium az írói-költői életművek feldolgozása számára, melyeket a könyv mint médium és mint szöveghordozó fizikai tulajdonságainál fogva nem tud biztosítani.

Bernard Cerquiglini 1989-ben publikált tanulmányában,⁴ *A variáns dicséretében* úgy véli, hogy a genetikus kritika a nyomtatott szöveget megelőző és annak keletkezési fázisait jelentő előszövegek leírásával ugyan szemléltetni képes az írás instabilitását, mégis a modern szövegre koncentrálni, mely a könyvnyomtatás után alakult ki, és a jogi keretek között is kodifikált szerzőség védjegye alatt végül is a nyomtatott szövegben stabilizálja a variánsaiban instabil írást. Ezzel szemben a középkori újlatin irodalom filológusaként Cerquiglini úgy látja, a másolatainak sokféleségében létező középkori szöveg alapvető jellemzője a variancia, és szemben a korábbi (lachmannista, bédier-ista) szövegkiadói gyakorlattal a textológusnak szakítania kell a stabil szöveg pozitívista illúziójával és az eredeti archetípusból levezetett genetikus leszármazási rajzok „megható fikciójával”.⁵ A középkori szöveg változékonyságát Cerquiglini szerint leginkább a számítógép tudja szimulálni, ugyanis a nyomtatott szöveg kétdimenziós és zárt szerkezetét felváltja a variánsok „képernyőszerű” elrendezése, mely egyrészt dialogikus, tehát lehetővé teszi a gép és a felhasználó folyamatos interakcióját, másrészt többdimenziós, azaz a különböző szövegegységeket egy időben több ablakban tudja megnyitni, s ezáltal lehetővé teszi azok összekapcsolt vizsgálatát. Cerquiglini pontosan látja azt a lehetőséget is, amit a számítógép nyújt a tudományos szövegkiadás számára a nyelvin túli „járulékos” információk (konkordancia, előfordulási gyakoriság, rímtáblázat, kodikológiai és paleográfiai információk stb.) közlésével. „Az informatika nemcsak kielégíti a teljesség szükségletét – állapítja meg Cerquiglini –, hanem a szöveg megértését is támogatja. És ebben látjuk a legjelentősebb és legfontosabb technikai hozadékot: a számítógép segíthet a szöveg mozgásának megfigyelésében, láthatóvá téve a kapcsolatokat, melyeket a kiadó előkészített és sugalmazott.”⁶

Cerquiglini a számítógép jelentőségét a könyvnyomtatáshoz viszonyítja, s hasonlóan jelentős kultúraformáló hatást tulajdonít neki. Álláspontja azonban távol áll azoktól a kultúrpeszsimista jóslatoktól, melyek az új médium térhódításával az írás (és olvasás) háttérbe szorulását prognosztizálják:

⁴ Bernard CERQUIGLINI, *Éloge de la variante: Histoire critique de la philologie* (Paris: Ed. du Seuil, 1989). Magyarul: Bernard CERQUIGLINI, „A variáns dicsérete: A filológia kritikai története”, ford. KESZEG Anna, in *Metafilológia I: Szöveg, variáns, kommentár*, szerk. DÉRI Balázs, KELEMEN Pál, KRUPP József és TAMÁS Ábel, *Filológia* 2, 219–297 (Budapest: Ráció Kiadó, 2011).

⁵ CERQUIGLINI, „A variáns dicsérete...”, 256.

⁶ Uo., 296.

A technikai, majd kulturális elmozdulás, amelyben részt veszünk, nem kérdőjelezi meg az írásbeliséget, sőt éppen ellenkezőleg. Az informatika is az írásbeliség egy technikája. Egyfelől fantasztikus hatékonysággal tölti be az írás egyetemesen pozitív és fejlődéselvű funkcióit (letétbe helyezni, osztályozni, újra megtalálni és terjeszteni a tudást). Másfelől óriási mennyiségű írásos anyagot használ fel, hoz létre, raktároz és terjeszt el. Végül pedig az írás és olvasás új eljárásait vezeti be.⁷

Cerquiglini számára a számítógép nyújtotta „képernyőszerű” elrendezés a szöveg dialogikus és többdimenziós reprezentációs lehetőségeként elsősorban a középkori szöveg varianciájának bemutatása szempontjából vált fontossá, s éppen azért, mert olyan eszközt ad a textológus kezébe, melyet a hagyományos nyomtatott könyv nem tudott biztosítani számára. A szövegváltozatok „képernyőszerű” elrendezése azonban nemcsak a középkori szöveg varianciáját képes szemléltetni, hanem a modern művek keletkezésének fázisait – melyre éppen a genetikus kritika irányította a figyelmet – is reprezentálni tudja. Hasonlóan pragmatikus oldalról közelíti meg a problémát 1995-ös *The Rationale of Hypertext* című tanulmányában Jerome McGann is.⁸ McGann úgy látja, hogy „az irodalom, amit (mind a mai napig) öröklünk és átörökítünk, most és mindörökké könyvformájú lesz.”⁹ Másrészt azonban azt is érzékeli, hogy éppen a kritikai szövegkiadás számára kínálja a számítógép olyan eszközöket, amelyek segítségével a könyvformátum korábban átléphetetlennek vélt korlátain mégis sikerül átjutni. McGann alapképlete szerint a hagyományos kritikai kiadás *könyv segítségével tanulmányoz egy másik könyvet*, a szövegkiadás változatai a facsimilétől a kritikai kiadásig tehát azt a célt szolgálják, hogy a rendelkezésükre álló eszközökkel (jegyzetapparátussal) megvilágítsák egy másik szöveg jelentését. A probléma éppen a könyvformátum fizikai korlátaiban rejlik. A szöveg és szöveg közötti közvetítésen túl a nyomtatott kritikai kiadás nem tudja visszaadni a művekhez kapcsolódó nem nyelvi attribútumokat, mint például egy ballada vagy egy színdarab előadását, esetleg a szövegek fizikai jellemzőit.¹⁰ Az elektronikus szövegkiadás a nyomtatott kiadásnak a könyvformában rejlő két alapvető problémáját kezeli. Az olvasónak az elektronikus szöveg esetében nem kell bonyolult rövidítési rendszereket dekódolnia, melyek a genetikus kiadás egyik nehézségét jelentik, ezek a műveletek a képernyőn megjelenő elektronikus szövegben is jelen vannak,

⁷ Uo., 221.

⁸ Jerome McGANN, *The Rationale of Hypertext*, hozzáférés: 2018.12.17, <http://www2.iath.virginia.edu/public/jjm2f/rationale.html>. Magyarul: Jerome McGANN, „A hipertext alapjai”, ford. ÁRKOS Eszter, *Helikon* 50, 3. sz. (2004): 366–385.

⁹ McGANN, „A hipertext alapjai”, 367.

¹⁰ Uo., 368.

[u]gyanakkor a digitális kiadásban az olvasónak nemcsak, hogy nem kell megtanulnia, de még találkoznia sem kell az alapvető fontosságú műveletek végrehajtásához szükséges kódokkal (pl. oda-vissza mozogni egy vagy több könyv különböző részei között, elemzés céljából kereséseket és összehasonlításokat végezni). Ezek a műveletek a felhasználó utasítására, ám láthatatlanul, a háttérben történnek.¹¹

Másrészt nem kell az eredeti szöveg radikálisan átalakított (azaz jórészt csak a nyelvi kódot megőrző) változatát olvasni.

Ma a digitális bölcsészetben alkalmazott legelterjedtebb szabvány a TEI XML. Maróthy Szilvia Lou Burnard nyomán¹² a TEI XML három alapelvét emeli ki: (1) a TEI XML a szöveg jelentésére fókuszál és nem a megjelenítésére; (2) szoftverfüggetlen; (3) a tudományos közösség számára készült, mely folyamatosan gondoskodik a szabvány fejlesztéséről.

A TEI XML kódolás legfőbb célja tehát a tudományos adat hosszú távú megőrzése, olvashatóságának fenntartása és összehasonlíthatóságának biztosítása. Ebből következik, hogy a nagyobb szövegtudományok feldolgozó tudományos kiadások inkább működnek szakadatbázisként, mint egyszerű olvasókönyvként. A szövegek egységes elvek szerinti feldolgozása ugyanis lehetővé teszi azok sok szempontú visszakereshetőségét.¹³

A hipermediális környezetben működő archívumra legkézenfekvőbben a McGann főszerkesztésében készült Rossetti Archive¹⁴ projektre hivatkozhatunk, mely a 19. századi olasz származású angol író, költő és festő életművét feldolgozó online archívum. A Rossetti Archive fejlesztése még 1993-ban kezdődött, s a Rossetti-életművet több részletben feltöltve egészen 2008-ig tartott. A Rossetti Archive a NINES¹⁵ (Networked Infrastructure for Nineteenth-century Electronic Scholarship), a 19. századi angol és amerikai irodalom és kultúra anyagát nyílt hozzáféréssel, valamint lektorált tartalommal közzé tevő online projekt első próbaverziója volt. A NINES archívumaiban a szöveges és a képi anyag megjelenítését, valamint elemzését nyílt forráskódú szoftverek (Collex, IVANHOE, Juxta) segítségével oldják meg, a Rossetti Archive-ban az egyes műveknek megtaláljuk a digitalizált fotóját, festmények és kéziratok esetében az eredeti másolatát, irodalmi művek esetében az első közlés digitális kópiáját. Minden anyagot XML kódolás-

¹¹ Uo., 379–380.

¹² LOU BURNARD, „Introduction”, in LOU BURNARD, *What Is the Text Encoding Initiative?: How to Add Intelligent Markup to Digital Resources*, Encyclopédie Numérique (Marseille: OpenEdition Press, 2014), <http://books.openedition.org/oejp/679>.

¹³ MARÓTHY Szilvia, „Tudományos szövegkiadások a hálózaton: Áttekintés”, *Irodalomtörténeti Közlemények* 122, 5. sz. (2018): 617–633, 624–625.

¹⁴ Rossetti Archive, hozzáférés: 2018.08.09, <http://www.rossettiarchive.org/>.

¹⁵ NINES, hozzáférés: 2018.08.09, <http://www.nines.org/about/>.

sal átírtak, ezáltal az oldalba beépített keresőmotor segítségével az archívum anyagában strukturált keresés végezhető.

2018-as tanulmányában Maróthy Szilvia a fogalomtisztázás és a rendszerezés igényével tekintette át a hazai tudományos hálózati szövegkiadást, tanulmányában közölt listáján 52 tételt sorol fel.¹⁶ Az alábbiakban a teljesség igénye nélkül néhány olyan nagyobb volumenű vállalkozást szeretnék ismertetni, melyek több tétellel is gazdagították a listában összefoglalt magyarországi hálózati szövegkiadásokat.

A BIÖP HÁLÓZATI KRITIKAI KIADÁSAI

A számítógép bölcsészettudományi alkalmazásának magyarországi története több évtizedes múltra tekint vissza, már a *Helikon* 1969. évi 3–4. (vegyes) számának tematikája *A számítógép és a humán tudományok* volt. A hetvenes években a szegedi egyetemen indult a *Répertoire de la poésie hongroise ancienne*, a régi magyar irodalom francia nyelvű repertóriumának összeállítása. Az RPHA nemzetközi beágyazottságú projekt volt, a Párizsban működő Polivanov Kör *Ad Corpus Poeticarum* projektjéhez csatlakozott, melynek célja a világ valamennyi verselési rendszerének leírása volt.¹⁷ Az 1601-ig született versek repertóriumának összeállítása számítógépes környezetben zajlott, kezdetben peremlyukkártyás megoldással, majd nagyszámítógépen, később személyi számítógépen, végül az összeállítók hálózati kiadásként publikálták.¹⁸ A Szegedi Számítógépes Munkacsoport jogutódja¹⁹ a Horváth Iván által vezetett, ELTE BTK keretei között működött Bölcsészettudományi Informatika Önálló Program volt. „A BIÖP tudományos műhelyében a tartalomgyártás új technológiáit kutatták, és a gyakorlatban ki is próbálták. A program keretében az online irodalom tartalmi és technikai kérdéseivel foglalkoztak, illetve hálózati kritikai kiadásokat hoztak létre.”²⁰ A BIÖP weboldaláról²¹ több hálózati kritikai kiadás is elérhető, melyek nem nyomtatott könyvek digitális verziói (tehát – például – nem pusztán fotómásolatokat tesznek közzé PDF-ben), hanem eleve hálózati kiadásként (*born digital*)²² készültek, és megoldásokat keresnek azokra a problémákra (például a jegyzetapparátus kezelésére), melyek a ha-

¹⁶ MARÓTHY, „Tudományos szövegkiadások a hálózaton...”, 630–633.

¹⁷ TÓTH Tünde, „Irodalomtörténészek a bábeli könyvtárból: A régi magyar vers repertórium”, *Irodalomtörténeti Közlemények* 101, 1–2. sz. (1997): 133–145, 133–134.

¹⁸ Használati útmutató a régi magyar vers számítógépes repertóriumához, hozzáférés: 2018.08.14, <http://magyar-irodalom.elte.hu/repertorium/dokument/index.html>. A projekt legújabb, 2002-ben lezárt verziója elérhető: „Ad Corpus Poeticarum”, hozzáférés: 2018.08.14, <http://rpha.elte.hu/>.

¹⁹ Szűts Zoltán, „A magyar online irodalom publikálásának és recepciójának rövid története”, *Médiakutató* 17, 3–4. sz. (2016): 47–54, 53.

²⁰ Uo.

²¹ BIÖP – ELTE BTK Bölcsészettudományi Informatika Önálló Program, hozzáférés: 2018.08.14, <http://magyar-irodalom.elte.hu/>.

²² MARÓTHY, „Tudományos szövegkiadások a hálózaton...”, 626.

gyománys kritikai kiadások számára nehézséget jelentettek.²³ Készült olyan hálózati kritikai kiadás, mely megvalósította a különböző variánsok osztott képernyős szinoptikus olvasását (például az Orlovsky Géza által közzétett *Adriai tengerek syrenája*²⁴), vagy a betűhív átirat mellett publikálták az eredeti korpusz facsimiléjét is (*Wathay Ferenc összes művei*²⁵).

A többségében régi magyar irodalmi szöveganyag mellett két 20. századi hálózati kritikai kiadást is publikál az oldal: a Horváth Iván által főszerkesztett *József Attila összes tanulmánya és cikke* lezárt 1.0 és 2.0 verzióját,²⁶ és Kosztolányi Édes Anna című regényének demó verzióját.²⁷ A József Attila-projektnek két filológiai kutatómunka futott párhuzamosan, az egyik egy papíralapú kritikai kiadást célzott meg, hátterét a 2003-ban az OTKA által támogatott kutatás adta, melynek vezetője Tverdota György volt. Ő készítette a tárgyi jegyzeteket és a szövegmagyarázatokat, a szöveggörgető munkákat pedig Horváth Iván kutatócsoportja végezte. A jelenleg elérhető hálózati kiadás az 1999-es „kísérleti képzeteskönyv” továbbfejlesztett változata.

Horváth Iván a *József Attila összes tanulmánya és cikke* hálózati kritikai kiadásnak a metodikáját a genetikus kritika és a hagyományos kiadás kettősségében határozza meg:

A kiadás annyiban tér el a szokásos genetikus kiadásoktól, hogy azok nem a szöveget, hanem a kéziratot akarják nyújtani az olvasónak, emez viszont – a régimódi kiadásokhoz hasonlóan – a kézirat esetlegességeit zajnak tekinti, amelyektől meg kell szabadítani a szöveget. A szöveg megalkotásának rétegeit viszont tökéletesen elkülöníti egymástól, mindegyik réteget önmagában is, zavartalanul olvashatóvá teszi, grafikusán ábrázolni tudja a rétegek átalakulását egymásba, sőt – különleges olvasói igények esetén – olyan rétegeket is láthatóvá tesz, amelyeknek nem maradt filológiai nyomuk, de amelyek a megalkotás folyamata során átmenetileg léteztek a költő képzetében.²⁸

²³ Az ELTE Magyar Irodalomtörténeti Intézetének Reneszánszkutatások Posztgraduális Központja (Centre des Hautes Études de la Renaissance, CHER) és a BIÖP által kiadott hálózati kritikai kiadásokról bővebben: Tóth Tünde, „Online kritikai szövegkiadás Magyarországon az ezredfordulón”, *Helikon* 50, 3. sz. (2004): 417–442, 431–433.

²⁴ ORLOVSKY Géza, szerk., *Adriai tengerek syrenája, gróf Zrínyi Miklós* (Budapest: ELTE BTK Régi Magyar Irodalomtörténeti Tanszék – Gépeskönyv, 1998), <http://magyar-irodalom.elte.hu/gepesk/zrinyi/>.

²⁵ Tóth Tünde és Tóth Csilla, szerk., *Wathay Ferenc összes műve: hálózati kritikai kiadás* (Budapest–Zalaegerszeg: Gépeskönyv, 1999), <http://magyar-irodalom.elte.hu/gepesk/wathay/>.

²⁶ HORVÁTH Iván, DEVESCOVI Balázs és GOLDEN Dániel, szerk., *József Attila összes tanulmánya és cikke, v1.0 és v2.0, Hálózati kritikai kiadássorozat* (Gépeskönyv, 1999), <http://magyar-irodalom.elte.hu/ja/>.

²⁷ VERES András, szerk., *Kosztolányi Dezső: Édes Anna: hálózati kritikai kiadás, Hálózati kritikai kiadássorozat* (Gépeskönyv, 2001), <http://magyar-irodalom.elte.hu/gepesk/kd/edessanna/>.

²⁸ HORVÁTH Iván, *Gépeskönyv, Opus Irodalomelméleti Tanulmányok 9* (Budapest: Balassi Kiadó, 2006), 343.

A tartalomjegyzék keletkezési és alfabetikus rendben közli a címeket, a nagyobb értekezőpróza írások (*Ihlet és nemzet, Irodalom és szocializmus, Hegel, Marx, Freud* stb.) variánsait az adott mű címe alatt találjuk. Az egyes szövegek sorszámozott átiratában jegyzetek jelzik a szövegkritikai munka helyét, a jegyzetszámokra kattintva felugró ablakban kapjuk meg a vonatkozó textológiai információkat (lásd az 1. ábrát).

Az olvasó a szöveg melletti oldalsávban olvashatja a kézirat vagy a nyomtatott forrás filológiai adatait. A megfelelő ikonra kattintva megtaláljuk a forrás fotokópiáját, illetve a különböző variánsok összehasonlítása érdekében különböző módú osztott képernyőn a szöveg párhuzamos olvasására is lehetőséget ad a hálózati kiadás (lásd a 2. ábrát).

2012-ben Horváth Iván és Fuchs Anna szerkesztésében elkészült a *József Attila összes tanulmánya és cikke, 1930–1937* hálózati kiadása is,²⁹ mégis – és ez a digitális szövegkiadás mai magyarországi tudományos életben belüli presztízse szempontjából is rendkívül tanulságos – az 1930 után keletkezett értekező próza írások újabb szövegkritikai revíziója nem webes környezetben történt meg, a József Attila Társaság kutatócsoportja által végzett újabb kutatások eredményét ugyanis nyomtatott kritikai kiadásban tették közzé.³⁰ Az új nyomtatott kritikai kiadás szerkesztői előszava pusztán a szövegkritikai problémákkal indokolja az új kiadás szükségességét, és szó nélkül hagyja a médium(vissza)váltás problematikáját:

Horváth Iván és munkacsoportja úttörő munkát végzett a szövegek lejegyzésében (továbbá digitális közzétételében és népszerűsítésében); számos eddig publikálatlan kéziratot elsőként hozott nyilvánosságra és megpróbált rendet teremteni a kronológiában is. Kritikai kiadásunk kezdetekben a rendelkezésünkre bocsátott, általuk készített szövegközlésekre kívánt támaszkodni, de számos olyan probléma merült fel, melyekre csak az eredeti források tanulmányozása adhatott magyarázatot. Ennek során be kellett látnunk, hogy az interneten is olvasható – korábbi időpontokból származó előzményekre épülő, önkorrektív is felmutató – 2012-es hálózati kiadás (JAHSZK 2012) közléseiben nem csak átírási hibákat, téves olvasatokat kell javítani. Esetenként összetartozó szövegeket szakítottak szét és különállóakat tereltek egybe, szerzői szándékkal nem igazolt kompilált változatokat hoztak létre, a kéziratok egy részét nem vették figyelembe, filológiai gyakorlatuk pedig túlbonyolított, ugyanakkor nem kellően informatív. Ezért volt szükség a szövegkritikai módszertan átalakítására, a keletkezési időrend és az eredeti források felülvizsgálatára.³¹

²⁹ HORVÁTH IVÁN és FUCHS ANNA, szerk., *József Attila összes tanulmánya és cikke: Szövegek 1930–1937, v1.0* (Eötvös Loránd Tudományegyetem, 2012), <http://textus.elte.hu/ja/>.

³⁰ JÓZSEF ATTILA, *Összes tanulmánya és cikke, 1930–1937*, szerk. TVERDOTA György, VERES András, SÁRKÖZI Éva, AGÁRDI Péter és BÉRES Judit, Kritikai kiadás (Budapest: József Attila Társaság–L'Harmattan Kiadó, 2018).

³¹ Uo., 9.

Előszó Tartalomjegyzék Sűgő Névjegy

[7] [A művészi tényről...]

I. Mí a műalkotás?

A képzés története. A művészi tényről való történeti felfogások egymás cáfolásai. Hol az egyszerű intuícióval azonosították, hol az "összetett" valódi képzetet, hol pedig a tárgyiasított érzékeltetést szemben a képzett képzetével. Bevonták az esztétika tárgykörébe ha nem is éppen az emberi szerzevet zörögést, hanem az indulat megnyilvánulását, minthogy – ha a művészet intuíció – a műalkotás csupán mennyiségileg különbözik tőlük, ami tehát a minőségek tudományára nézve közömbös. Így azonban – amit észre sem vettek – kivonták az esztétika köréből annak eldönését, hogy mi teszi a² kategóriáuk különbségét a műalkotás miné³ alak, és a nem műalkotás mint alak között, hogy miért műalkotás Diane Divina Comediója és miért nem az tetzem a következő kifejezés: ez a toll. Másrésztől azonban még zavarosabb az az állítás, amely szerint a művészet az intuíció rejteje, mert ilyen nincs, ez a kapcsolás úres és hamis. És nyilvánvaló⁴ hogy ezeknek a tévedéseknek alapja abban a módszerben rejlik, ahogyan rejtejt, vagy kevésbé rejtejt pszichológizmussal próbálják megközelíteni a műalkotást, ahelyett hogy önmagában szemlélték volna az érték volna a spekuláció tárgyává, holott valaminek a történeti vizsgálata el nem vészt a fogalmi megismerésig, egyszerűen azért, mert így módon a dolog megismerése helyett legfőbb a dolog történetének megismeréséig jutunkunk el, ami azonban már nem az esztétikára tartozik, hiszen ebben a folyamatban maga az esztétika is szenvedélyes szerepet játszhat csupán.

Az "intuíció" meghatározás elégtelensége. Fogadjuk el, hogy a művészet intuíció. E meghatározás semmit mondó voltát azonnal látjuk, ha teljesen azonosítani akarván véde, megfordítjuk a tételt, hogy tehát az intuíció művészet. Kétségetelen, hogy amilyen szabadon állíthatom az előbbi, amilyen szabadon állíthatom az utóbbi tételt, hiszen a poszta

A közlés alapja:
Gépirat, 1 f., 340 x 21.
A gépirat a lap alján megstábed.
Súlyok NIKLÓD
tulajdon, letét a PTH
Készírtárterében,
álatalunk adott
jelzete: 22, az 1.
Ídegen kéz
tételiszámában: 22.
Íróféfé: 1. Javítási szakasz:
szakasz: cruza.
Aláírta:lam.

Jegyzet - Google ...

about:blank

[Előbb:] történeti [végül: főszöveg]

A közlés alapja:
Gépirat, 1 f., 340 x 21.
A gépirat a lap alján megstábed.
Súlyok NIKLÓD
tulajdon, letét a PTH
Készírtárterében,
álatalunk adott
jelzete: 22, az 1.
Ídegen kéz
tételiszámában: 22.
Íróféfé: 1. Javítási szakasz:
szakasz: cruza.
Aláírta:lam.

1. ábra

Előszó Tartalomjegyzék Sűgő Névjegy

[8] A kompozícióról [1]

Köz hely, hogy az esztétika, vagy ha úgy tetszik, a szépről szóló tan éppolyan önálló tudomány, mint tetsem a logika,² mégis gondnokság alatt áll, éspedig különösképpen a pszichológia mit rendre. Es bír kétfélefen, hogy a műalkotás adta hatás vizsgálata is értekes és rendszerbe foglalható, azonban az idevágó kutatás eredménye legalább is problematikus, amíg magának a hatás³ forrásának műbenléte tisztázatlan, amíg vonatkozó kérdéseire határozott felelet nem adánk. Végünk a hatásától megfosztott művészi tényről normák s e művészi tény állításuk vissza a maga valóságában a pszichika tények köze: azonnal belátjuk, hogy míg ilyen normához semmiféle pszichológizmus nem juttathat, addig az a pszichológia számára a további specializált vizsgálódások terén csak üdvös és hasznos lehet. Emnyiből is látható, hogy a pszichológizmus a maga elvakultságában megoldná utbagaztatóját, ha ilyenről egyáltalában szó is lehetne. Beismerendően különböz,

A közlés alapja:
Gépirat, 1 f., 328 x 249 mm.
Számolás géppel: 3. A felbontás:
#1157b, az 1. ídegen kéz tételiszámában: 27.
Íróféfé: 2. Javítási szakasz: cruza; 1. ídegen kéz.
Aláírta:lam.
Az 1. lap tetején az 1. ídegen kéz:

Jegyzet - Google ...

about:blank

[Előbb:] történeti [végül: főszöveg]

A közlés alapja:
Gépirat, 1 f., 328 x 249 mm.
Számolás géppel: 3. A felbontás:
#1157b, az 1. ídegen kéz tételiszámában: 27.
Íróféfé: 2. Javítási szakasz: cruza; 1. ídegen kéz.
Aláírta:lam.
Az 1. lap tetején az 1. ídegen kéz:

2. ábra



3. ábra

A másik 20. századi mű Kosztolányi *Édes Anna* című regényének demója, melyet Veres András rendezett sajtó alá, ő írta a szövegkritikai és magyarázó jegyzeteket is.³² A „demó verzió”-megnevezés arra utal, hogy ez a hálózati kiadás nem tekinthető véglegesnek, a kiadó 1.0 verzióban ígéri a teljes, facsimilében közölt kéziratos és nyomtatott források digitális filológiai feldolgozását. Az *Édes Anna* hálózati kiadásában a kritikailag megállapított szövegbe hiperhivatkozásként szinkódokkal ellátott tárgyi és szómagyarázó, valamint szövegkritikai jegyzetek épülnek be, ezekre kattintva a jobb oldalon keretben jelenik meg a jegyzet szövege, a sáv alján megjelenő ikonra kattintva pedig a felhasználó eldöntheti, hogy melyik szövegvariánst (vagy esetleg mindhármát) kívánja megtekinteni (lásd a 3. ábrát).

A fejléc „FORRASOK” gombjára kattintva az oldal felajánlja a regény kéziratának, *Nyugatban* közölt változatának és első, 1926-os kiadásának megtekintéseit. Az autográf és a regény Genius Kiadónál megjelent kiadásának beszkenelt (fekete-fehér) másolatai a böngészőben megnyíló új lapon érhetők el. A BIÖP által kiadott *Édes Anna* végül demó maradt, a regényt a Kosztolányi Kutatócsoport keretében Veres András szerkesztette,³³ a kritikai kiadás digitális filológiai feldolgozása pedig a DigiPhil-projekt keretében került publikálására.³⁴

³² VERES, *Kosztolányi Dezső: Édes Anna: hálózati kritikai kiadás.*

³³ KOSZTOLÁNYI Dezső, *Édes Anna*, szerk. VERES András, s. a. r. PARÁDI Andrea, JÓZAN Ildikó, VERES András, SÁRKÖZI Éva és LIPA Tímea (Pozsony: Kalligram Kiadó, 2010).

³⁴ KOSZTOLÁNYI Dezső, *Édes Anna*, szerk. VERES András, PARÁDI Andrea, XML szerkesztő: MIHÁLY Eszter (Budapest: DigiPhil, 2015), <http://digiphil.hu/context:edesanna>.

AZ MTA–DE KLASSZIKUS MAGYAR IRODALMI TEXTOLÓGIAI KUTATÓCSOPORT
ELEKTRONIKUS KRITIKAI KIADÁSAI

Az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport eredeti feladata a Csokonai kritikai kiadás elkészítése volt, intézmények közötti együttműködés keretében ebből nőtt ki egy olyan textológiai program, mely a nyomtatott kiadványok elkészítése mellett jelenleg tizenkét szövegtörzset digitális filológiai feldolgozását teszi közzé.³⁵ A kutatócsoport digitális filológiai projektjei eredetileg azokra a problémákra kívántak megoldást találni, melyek a Csokonai- és a Kazinczy-életmű hagyományos textológiai feldolgozása és genetikai kritikai kiadása során felmerültek. Ezen kívül a kutatócsoport olyan fejlesztéseket is végrehajtott, melyek révén a rendelkezésre álló anyag feldolgozását, digitális filológiai reprezentációját és kutathatóságát nagy mértékben segíteni tudják: „Az elektronikus kiadás az a közeg, amelyben a genetikai szemlélet igazán otthonos lehet, s itt kell sor kerüljön a könyv formájú kritikai kiadás lezárulta óta felvetődött, imént jelzett filológiai problémák megoldására is, beleértve a költői életmű új kronológiájának megteremtését” – írja Csokonai életművének kronológiai rendjét tárgyalva a kutatócsoport-vezető, Debreczeni Attila.³⁶

Az elektronikus kiadás lehetőséget ad arra, hogy a hagyományos textológia ultima manus (vagy ultima editio) elve alapján kritikailag megállapított „ideális” (azaz a szerző végső szándékát tükröző) mű helyett a szövegforrások kerüljenek előtérbe. Az autográfok mellett azonban a Csokonai-életmű filológusa nem mondhat le a nem saját kezű másolatok vizsgálatáról sem, így az életmű digitális filológiai feldolgozása alkalmat ad a befogadás folyamatát bemutató variogenetikai szövegtér szemléltetésére is.³⁷ A Csokonai elektronikus kritikai kiadás³⁸ is megtartja az ún. „olvasószöveget”, mely az emendált és „olvasóbarát” állapotban közölt szöveget jelenti, emellett azonban a digitális sajtó alá rendezők közzé teszik a „kritikai szöveget” is, tehát a szövegforrás betűhű átiratát és szövegvizsgálati feldolgozását. Felugró ablakokban teszik láthatóvá a szerzői javításokat, valamint a sajtó alá rendező emendációit. A „genetikai szöveg” különböző variánsokat tesz egymás mellett láthatóvá és ezáltal összevethetővé (lásd a 4. ábrát).

A Debreczeni Attila által kidolgozott terminológia szerint ez adja az ún. szövegidentitás megtestesülését,³⁹ azt az egységet, „amely a különböző szövegváltozatokban szövegszerűen megnyilatkozik, az eltérések ellenére is. Lényege valójában a szöveg alakulási folyamata, amely a változatok összességében testesül meg,

³⁵ BODROGI Ferenc Máté, „Elektronikus kritikai kiadások a klasszikus magyar irodalomban”, *Magyar Tudomány* 177, 11. sz. (2016): 1292–1296.

³⁶ DEBRECZENI Attila, *Csokonai költői életművének kronológiai rendje* (Budapest: Akadémiai Kiadó–Debreceni Egyetemi Kiadó, 2012), 17.

³⁷ Uo., 23.

³⁸ DEBRECZENI Attila és TÓTH Barna, szerk., *Csokonai Vitéz Mihály összes művei: Elektronikus kritikai kiadás* (MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport, 2016), https://doi.org/10.5484/csokonai_vitez_mihaly_osszes_muvei.

³⁹ Vö. BODROGI, „Elektronikus kritikai kiadások a klasszikus magyar irodalomban”, 1294.

Egy Tulipánthoz Rózsá-versek. 12. (11a.) 1797 A		Egy Tulipánthoz Ódák. 31. 1802-1804/1805 A	Tartózkodó Kérelm Lilla. 44. 1802-1803/1805 B
A hatalmas szerelemnek Megemésztő tüze bánt Te vagy orvossa sebémmek Gyönyörű kis tulipánt!	A HATALMAS Szerelemnek: Megemésztő tüze bánt: Te vagy orvossa sebémmek, Gyönyörű kis Tulipánt!	A HATALMAS Szerelemnek Megemésztő tüze bánt. Te lehetsz írja sebémmek Gyönyörű kis Tulipánt!	
Szemeid szép ragyogása Gyönyörű hajnali tűz, Ajakid' harmatozása Sok ezer gondot el' üz.	Szemeid' szép ragyogása Lobogó hajnali tűz; Ajakid' harmatozása Sok ezer gondot' angyaliszókkal Sh., em.	Szemeid' szép ragyogása Eleven hajnali tűz, Ajakid' harmatozása Sok ezer gondot el'üz.	
Teljesíted angyali szókkal Szeretéd a' mire kért Ezer ambrosia csókkal Fizetek tsokjaidért.	Teljesíted angyali szókkal. Szeretéd a' mire kért: Ezer ambrosia csókkal Fizetek* csókjaidért.	Teljesíted angyali szókkal. Szeretéd a' mire kért: Ezer ambrosia csókkal Fizetek válaszdért.	

4. ábra

beleértve a (filológiai vagy logikailag) kikövetkeztethető, de fizikailag nem létezőket is.”⁴⁰ Minden szövegforráshoz kétféle navigátor kapcsolódik. Az ún. szövegforrás-navigátor táblázatos formában tartalmazza a forrásra vonatkozó legfontosabb adatokat. Ebből a táblázatból lehet átlinkelni a társszövegforrásokra, a többi szövegváltozatra, az autorizált másolatokra és a gyűjteményes kötetekre, valamint a genetikus szövegre és a szövegidentitás-generátorra, mely a források egymáshoz való viszonyait mellérendelten ábrázolja.⁴¹

Az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport elektronikus szövegkiadásai kihasználják a web adta hipertextuális és hipermediális lehetőségeket (megtaláljuk itt például Hovánszki Mária gondozásában Csokonai énekelt költészetének hanganyagát),⁴² az oldal keresőmotorja segítségével a felhasználó összetett kereséseket végezhet a kiadások háttérét adó adatbázisban. A kutatócsoport elektronikus szövegkiadásainak jelentős részét teszik ki az írói levelezések. A *Magyar írók levelezése*⁴³ mellett *Kazinczy Ferenc összes művei*⁴⁴ között megtaláljuk az író, valamint önálló kiadásban Berzsenyi Dániel⁴⁵ és a Ráday-csa-

⁴⁰ DEBRECZENI, *Csokonai költői életművének kronológiai rendje*, 19.

⁴¹ BODROGI, „Elektronikus kritikai kiadások...”, 1294.

⁴² D. HOVÁNSZKI Mária, szerk., *Csokonai Vitéz Mihály énekelt költészete (elektronikus kritikai kiadás)*, DVD-ROM (Debreceni Egyetemi Kiadó, 2009).

⁴³ DEBRECZENI Attila és ORBÁN László, szerk., *Magyar írók levelezése: Elektronikus kritikai kiadás* (MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport, 2016), https://doi.org/10.5484/Magyar_irok_levelezese.

⁴⁴ BODROGI Ferenc Máté, ORBÁN László és DEBRECZENI Attila, szerk., *Kazinczy Ferenc összes művei: Elektronikus kritikai kiadás* (MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport, 2016), https://doi.org/10.5484/Kazinczy_Ferenc_muvei.

⁴⁵ DEBRECZENI Attila, TÓTH Barna és FÓRIZS Gergely, szerk., *Berzsenyi Dániel levelezése* (MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport, 2017), https://doi.org/10.5484/Berzsenyi_Daniel_levelezese.

lád⁴⁶ levelezését. Bodrogi leírása szerint: amíg az írói portálok szövegforrás-központúak, addig a levelezéskiadások kapcsolatközpontúak, ami azt is jelenti, hogy a kiadások révén felvázolhatóvá válik a korszak értelmiségi kapcsolathálója. A személyes kapcsolatrendszer a levelek keletkezési helyének topográfiai rajza egészíti ki, mely pontosan rávilágít a kapcsolati hálók térbeli eloszlására is.⁴⁷

A DIGIPHIL⁴⁸

A DigiPhil (A magyar irodalomtudomány filológiai portálja) az előbb ismertetthez képest más, egyszerre szűkebb és szélesebb profillal működik. A DigiPhil projekt keretében jelenleg elérhető a Kosztolányi-⁴⁹ és a Mikszáth-életmű⁵⁰ kritikai kiadásának néhány kötete és Móricz Zsigmond világháborús naplói,⁵¹ illetve levelezésének egy része,⁵² valamint Kassák avantgárd mozgalmanak dokumentumai,⁵³ továbbá Arany János születésének bicentenáriuma alkalmából elkészült az Arany-kritikai kiadások digitalizációja is.⁵⁴ Találunk az oldalon a metafilológia tárgykörébe tartozó, eredetileg nyomtatásban megjelent fordításkötetet,⁵⁵ továbbá Kassák és az avantgárd mozgalom történetét feldolgozó tanulmánykötetet.⁵⁶ A DigiPhilről írt tanulmányában a projektvezető, Palkó Gábor a projekt négy funkcióját határozza meg, melyek közül pusztán az egyik a tudományos szöve-

⁴⁶ DEBRECZENI Attila és VEISZ Bettina, szerk., *A Ráday-család levelezése: Elektronikus kritikai kiadás* (MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport, 2016), https://doi.org/10.5484/A_Radai-csalad_levelezese.

⁴⁷ BODROGI, „Elektronikus kritikai kiadások...”, 1295.

⁴⁸ DigiPhil: A magyar irodalomtudomány filológiai portálja, hozzáférés: 2018.08.27, <http://www.digiphil.hu>.

⁴⁹ PARÁDI Andrea és VERES András, szerk., *Kosztolányi Dezső: Édes Anna* (Petőfi Irodalmi Múzeum, 2015), <http://digiphil.hu/context:edesanna>; VERES András és PALKÓ Gábor, szerk., *Kosztolányi Dezső: Esti Kornél* (Budapest: Petőfi Irodalmi Múzeum, 2016), <http://digiphil.hu/context:estikornel>; BENGÍ László és MIHÁLY Eszter, szerk., *Kosztolányi Dezső: Aranysárkány* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:aranysarkany>.

⁵⁰ LABÁDI Gergely, szerk., *Mikszáth Kálmán összes művei*, 39 (Budapest: Petőfi Irodalmi Múzeum, 2015), <http://digiphil.hu/context:mkk39>; [s. n.], szerk., *Mikszáth Kálmán összes művei*, 42 (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:mkk42>.

⁵¹ CSÉVE Anna, szerk., *Móricz Zsigmond: Tükör* (Budapest: DigiPhil, 2018), <http://digiphil.hu/context:tukor>.

⁵² CSÉVE Anna, szerk., *Móricz Zsigmond levelezése (1892–1913)* (Budapest: Petőfi Irodalmi Múzeum, 2019), <http://digiphil.hu/context:moriczlevelezes>.

⁵³ *A Tett folyóirat* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:atett>; *Kassák Lajos: A ló meghal a madarak kirepülnek: Fordítások* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:alomeghal>.

⁵⁴ PALKÓ Gábor és FELLEGI Zsófia, szerk., *Arany János összes művei* (Budapest: Magyar Tudományos Akadémia Bölcsészettudományi Kutatóközpont, Irodalomtudományi Intézet, Petőfi Irodalmi Múzeum, 2019), <http://digiphil.hu/context:ajom>.

⁵⁵ „Metafilológia I. (tanulmányok)”, DigiPhil, hozzáférés: 2020.02.26, <http://digiphil.hu/context:metafilologia1>.

⁵⁶ „Local Contexts / International Networks: Avant-Garde Journals in East-Central Europe”, DigiPhil, hozzáférés: 2018.08.22, <http://digiphil.hu/context:localcontexts>.

gek közzététele. Palkó emellett további funkcióként említi az adatgazdagítást és az adatbázisba rendezést, a más adatbázisokkal történő összekapcsolódást. Az adatgazdagítás és az adatbázisba rendezés célja a szövegek intelligens kereshetővé tétele. A DigiPhil által szolgáltatott szövegek szemantikus annotációja szintén a hatékony keresést kívánja szolgálni. Nagy adatbázisokkal való összekapcsolódás (aggregáció) „során a kisebb helyi gyűjtemények digitális anyagait, képeket, hangokat, videókat és szövegeket, illetve az ezeket leíró adatokat nagyobb, nemzeti vagy nemzetközi szolgáltatásokban összegyűjtik, feldolgozzák, gazdagítják, kereshetővé teszik.”⁵⁷ A DigiPhil az általa szolgáltatott adatokat az európai kulturális örökséget digitális formában őrző, több intézményi aggregátorral működő Europeanán osztja meg.⁵⁸ Mindez kisebb részben PDF-formátumú, nagyobbrészt TEI XML kódolású szöveges tartalmat szolgáltató⁵⁹ projekt keretében történik, mely korszerű informatikai háttérnek köszönhetően biztosítja a szövegek stabilitását az adatomódosulásokkal, illetve az adatvesztésekkel szemben, emellett „szabványosított és hierarchikusan strukturált, hálózatba rendezett metaadatokkal”⁶⁰ javítja a keresések találati hatékonyságát, a kutatók számára pontosan definiálható böngészések révén a szövegelemzés új lehetőségeit kínálja. Az adatgazdagítás részét képezi a könyvészeti metaadatok összekapcsolása más adatbázisokkal, a szerzői névről például tovább lehet lépni a Virtual International Authority File-ra, ahol a több nemzeti könyvtár szerzőségi adatainak kapcsolati hálóját lehet vizsgálni az adott szerzői névre vonatkozóan.

A DigiPhilen közzétett kritikai kiadások nem eleve digitálisan készült szövegek közlések, hanem úgynevezett digitális átültetések,⁶¹ azaz egy korábban kiadott könyvalapú kritikai kiadás átalakításai XML kiadássá. A szövegek variánsai esetében a DigiPhil is sokrétű adatszolgáltatást végez, kéziratok esetében közlik az autográfot őrző archívum nevét és a kézirat anyagi hordozójának jellemzőit. A Kosztolányi-művek jegyzetei például megadják az íróeszköz fajtáját, toll esetében a tinta színét, melynek jelentősége van a kézirat alaprétegeinek a megállapítása során. Az elektronikus kiadás lehetővé teszi az egyes művek kéziratosa, illetve nyomtatott variánsainak egymás mellett történő, Cerquiglini szavaival „képernyőszerű” elrendezését, a szövegkritikai jegyzet a szövegbe ágyazódik, az alaptól eltérő háttérszínnel kiemelve, a kurzort a jegyzetelt szövegrész mellett található csillag fölé helyezve válik láthatóvá (lásd az 5. ábrát).

⁵⁷ PALKÓ, „Mit jelent a digitális filológia a szemantikus web korában?”, 1320.

⁵⁸ Gyöngyösi János műveinek forráskiadását például a DigiPhil a Klasszikus Magyar Irodalmi Textológiai Kutatócsoporttal együtt végezte, ez szintén elérhető az Europeanában is.

⁵⁹ Sem az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport elektronikus kritikai kiadásainak, sem a DigiPhil szövegek közléseinek TEI XML forráskódjai nem nyílt hozzáférésűek, azaz nem elérhetők és nem letölthetők. Vö. MARÓTHY, „Tudományos szövegkiadások a hálózaton: Áttekintés”, 628.

⁶⁰ PALKÓ, „Mit jelent a digitális filológia a szemantikus web korában?”, 1319.

⁶¹ MARÓTHY, „Tudományos szövegkiadások a hálózaton...”, 626–627.

DIGIPHIL PROJ

Kosztolányi Dezső: Aranyársarkány

Összesen 4 verzió ▼ TELJES SZÖVEG « 1 ▼ 2 3 »

Verzió: K: Szerzői kézirat. A Magyar Tudományos

(Hány) [De ne hányd annyira a lábád.]

- Igen - bólintott Vili, mert tudta, hogy ez a hibája.

Liszner Vili nem (Hány)szépen futott. Ha stílust akart teremteni, mindig elrontotta a futamot. Csak (Hány) a marha-erejével bírta.

Egyik p (Hány)galány (Hány)panyó (Hány) - Vedd (Hány)Zöld tintával írt változtatás. Különb (Hány)is mára elég.

Azért még nehányszor lekuporodott, sprintelt, végigszaladta a pályát. Utána átengedte két barátjának, Packának, Huszár Bandinak, kik tizenkettő, tizenhárom alatt teszik meg s alapjában rokonszenves műkedvelőknek tekinthetők. Darabig nézte őket.

Verzió: stage1:

(Hány) [De ne hányd annyira a lábád.]

- Igen - bólintott Vili, mert tudta, hogy ez a hibája.

Liszner Vili nem (Hány)szépen futott. Ha stílust akart teremteni, mindig elrontotta a futamot. Csak (Hány) a marha-erejével bírta.

(Hány) nelt-e-a a] egy kabátot, vállára tette, (Hány) megfázol. Még hűvös van. Különb (Hány)is mára elég.

Azért még nehányszor lekuporodott, sprintelt, végigszaladta a pályát. Utána átengedte két barátjának, Packának, Huszár Bandinak, kik tizenkettő, tizenhárom alatt teszik meg s alapjában rokonszenves műkedvelőknek tekinthetők. Darabig nézte őket.

Verzió: stage2:

(Hány) [De ne hányd annyira a lábád.]

- Igen - bólintott Vili, mert tudta, hogy ez a hibája.

Liszner Vili nem (Hány)szépen futott. Ha stílust akart teremteni, mindig elrontotta a futamot. Csak (Hány) a marha-erejével bírta.

Egyik pajtása leemelt-e-a a galányabokorra[!] egy kabátot, vállára tette, panyókéra.

- Vedd föl - szólt - megfázol. Még hűvös van. Különb (Hány)is mára elég.

Azért még nehányszor lekuporodott, sprintelt, végigszaladta a pályát. Utána átengedte két barátjának, Packának, Huszár Bandinak, kik tizenkettő, tizenhárom alatt teszik meg s alapjában rokonszenves műkedvelőknek tekinthetők. Darabig nézte őket.

5. ábra

Ez a megoldás anélkül teszi lehetővé a variánsok egymásutánjában feltáruló szöveggenezis tanulmányozását, hogy a szövegekritikai jelek kvázi olvashatatlanná tennék a szöveget. A felhasználó maga döntheti el, hogy mely szövegváltozatokat kívánja együtt olvasni, emellett a fényképezőgépet szimbolizáló ikonra kattintva megtekintheti a kézirat facsimiléjét is.

Egy írói, költői életmű digitális filológiai feldolgozása során a hagyományos filológusi-textológusi kompetenciáknak újabbakkal kell gazdagodniuk, melyek elsősorban az új médium elméleti ismereteire és gyakorlati felhasználásának módzataira vonatkoznak. Mindezen túl minden hasonló projekt széles körű szakmai összefogást igényel, mely az adott életmű és irodalomtörténeti korszak legjelentősebb filológusait, a digitális bölcsészeket és az adott terület speciális igényeit is értő informatikusokat kell hogy magába foglalja. A terjedelmi korlátok leküzdése mellett említhetjük többek között a szövegváltozatok párhuzamos megjeleníthetőségét és olvashatóságát, a strukturált keresés révén a szövegek vizsgálatának új lehetőségeit.

VÁLOGATOTT BIBLIOGRÁFIA¹

NEMZETKÖZI SZAKIRODALOM

- ALBRECHT, Andrea, Toni BERNHART, Sandra RICHTER and Marcus WILLAND, Hg. *Quantitative Ansätze in den Literatur- und Geisteswissenschaften: Systematische und historische Perspektiven / Quantitative Methods in Literary Studies and the Humanities: Systematic and Historical Perspectives*. Berlin: De Gruyter, 2018. <https://doi.org/10.1515/9783110523300>.
- APOLLON, Daniel and Claire BELISLE. *Digital Critical Editions*. Champaign, Illinois: University of Illinois Press, 2014.
- AQUILINA, Mario. „The Work of the Literary Critic in the Age of Big Data”. *Interdisciplinary Literary Studies* 19, no. 4. (2017): 493–516. <https://doi.org/10.5325/intelitestud.19.4.0493>.
- BARTALESI, Valentina, Carlo MEGHINI, Daniele METILLI, Mirko TAVONI and Paola ANDRIANI. „A Web Application for Exploring Primary Sources: The DanteSources Case Study”. *Digital Scholarship in the Humanities* 33, no. 4. (2018): 705–723. <https://doi.org/10.1093/llc/fqy002>.
- BODE, Katherine. *A World of Fiction: Digital Collections and the Future of Literary History*. University of Michigan Press, 2018. <https://doi.org/10.3998/mpub.8784777>.
- BURNARD, LOU. *What Is the Text Encoding Initiative?: How to Add Intelligent Markup to Digital Resources*. Encyclopédie Numérique. Marseille: OpenEdition Press, 2014. <http://doi.org/10.4000/books.oep.426>.
- BUZZETTI, Dino. „Turing e l’elaborazione automatica del testo”. In *Convegno Per il centenario di Alan Turing fondatore dell’informatica*, 159–188. Contributi del Centro Linceo Interdisciplinare de Scienze Matematiche e Loro Applicazioni, 129. Roma: Accademia Nazionale dei Lincei; Scienze e Lettere, 2014.
- CIOTTI, Fabio, a cura di *Digital Humanities: Progetti italiani ed esperienze di convergenza multidisciplinare*. Collana Convegni 24. Roma: Sapienza Università Editrice, 2014. http://www.editricesapienza.it/sites/default/files/5094_Digital_Humanities.pdf

¹ A bibliográfia letölthető: <http://doi.org/10.5281/zenodo.3763817>. Köszönöm Golden Dániel, Kiss Margit, Kokas Károly, Parádi Andrea, Péter Róbert és Tubay Tiziano közreműködését a bibliográfia összeállításában.

- CIOTTI, Fabio e Gianfranco CRUPI, a cura di. *Dall'Informatica umanistica alle culture digitali: Atti del convegno di studi (Roma, 27–28 ottobre 2011) in memoria di Giuseppe Gigliozzi*. Quaderni Digilab 2. Roma: Casa Editrice Università La Sapienza, 2012. http://www.editricesapienza.it/sites/default/files/Quad_DigiLab_Informatica_Umanistica_Culture_Digitali.pdf.
- CRAIG, D. Hugh and Brett GREATLEY-HIRSCH. *Style, Computers, and Early Modern Drama: Beyond Authorship*. Cambridge: Cambridge University Press, 2017.
- DA, Nan Z. „The Computational Case against Computational Literary Studies”. *Critical Inquiry* 45, no. 3. (1 March 2019): 601–639. <https://doi.org/10.1086/702594>.
- DALEN-OSKAM, Karina VAN. „In Praise of the Variant Analysis Tool: A Computational Approach to Medieval Literature”. In *Texts, Transmissions, Receptions*, 35–54. *Modern Approaches to Narratives*. Brill, 2015. <https://www.jstor.org/stable/10.1163/j.ctt1w76wgh.7>.
- Debates in the Digital Humanities 2016*. Minneapolis–London: University of Minnesota Press, 2016.
- DOES, Jesse DE, Katrien DEPUYDT, Karina VAN DALEN-OSKAM and Maarten MARX. „Namespace: Named Entity Recognition from a Literary Perspective”. In *CLARIN in the Low Countries*, edited by Jan ODÍJK and Arjan VAN HESSEN, 361–370. Ubiquity Press, 2017. <https://www.jstor.org/stable/j.ctv3t5qjk.37>.
- DRISCOLL, Matthew James and Elena PIERAZZO, eds. *Digital Scholarly Editing: Theories and Practices*. Digital Humanities Series 4. Cambridge: Open Book Publishers, 2016. <https://doi.org/10.11647/OBP.0095>.
- EARHART, Amy E. *Traces of the Old, Uses of the New: The Emergence of Digital Literary Studies*. Editorial Theory and Literary Criticism. Ann Arbor, MI: University of Michigan Press, 2015. <https://doi.org/10.3998/etlc.13455322.0001.001>.
- EDER, Maciej, Jan RYBICKI and Mike KESTEMONT. „Stylometry with R: A Package for Computational Text Analysis”. *The R Journal* 8, no. 1. (2016): 107–121.
- EDMOND, Jennifer. *Digital Technology and the Practices of Humanities Research*. Cambridge: Open Book Publishers, 2020. <https://doi.org/10.11647/OBP.0192>.
- FIDLER, Masako and Václav CVRČEK, eds. *Taming the Corpus: From Inflection and Lexis to Interpretation*. Quantitative Methods in the Humanities and Social Sciences. Cham: Springer International Publishing, 2018. <http://doi.org/10.1007/978-3-319-98017-1>.

- FLANDERS, Julia and Fotis JANNIDIS, eds. *The Shape of Data in the Digital Humanities: Modeling Texts and Text-Based Resources*. London: Routledge, 2019.
- GABLER, Hans Walter. „Theorizing the Digital Scholarly Edition”. *Literature Compass* 7, no. 2. (2010): 43–56. <https://doi.org/10.1111/j.1741-4113.2009.00675.x>.
- HERRMANN, J. Berenike, Anne-Sophie BORIES, Francesca FRONTINI, Simone REBORA and Jan RYBICKI. „Response by the Special Interest Group on Digital Literary Stylistics to Nan Z. Da’s Study”. *Journal of Cultural Analytics*, 24 January 2020, 11827. <https://doi.org/10.22148/001c.11827>.
- ITALIA, Paola and Claudia BONSI. *Edizioni Critiche Digitali: Edizioni a confronto / Digital Critical Editions: Comparing Editions*. Collana Convegni 34. Roma: Sapienza Università Editrice, 2016. http://www.editricesapienza.it/sites/default/files/5369_Italia_Bonsi_EdizioniCriticheDigitali.pdf
- JOCKERS, Matthew L. *Macroanalysis, Digital Methods and Literary History*. Urbana; Chicago; Springfield: University of Illinois Press, 2013.
- KOHLE, Hubertus. „Die Geisteswissenschaften und das Digitale: Ein Quantensprung oder business as usual?”. München: Universitätsbibliothek der Ludwig-Maximilians-Universität, 2018. <https://doi.org/10.5282/UBM/EPUB.42414>.
- MAHMOUDI, M. R. and A. ABBASALIZADEH. „How Statistics and Text Mining Can Be Applied to Liter-Ary Studies?”. *Digital Scholarship in the Humanities* 34, no. 3. (2019): 536–541. <https://doi.org/10.1093/llc/fqy069>.
- MARINO, Mark C. *Critical Code Studies*. Cambridge, MA–London, UK: The MIT Press, 2020.
- MONELLA, Paolo. „L’edizione critica digitale: la critica del testo nella storia della tradizione”. *Storie e linguaggi. Rivista di studi umanistici* 5, no. 1. (2019): 61–82.
- MORETTI, Franco. „»Operationalizing«: Or, the Function of Measurement in Modern Literary Theory”. *Stanford Literary Lab Pamphlet*, 6 (December 2013). <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf>.
- NYHAN, Julianne and Andrew FLINN. *Computation and the Humanities: Towards an Oral History of Digital Humanities*. Springer Series on Cultural Computing. Cham: Springer International Publishing, 2016. <https://doi.org/10.1007/978-3-319-20170-2>.
- PERAZZINI, Federica. „The English Gothic Novel: Theories and Praxis of Computer-Based Macroanalysis in Literary Studies”. In *Big Data in the Arts and Humanities*, edited by Daniela CARLUCCI and Giovanni SCHIUMA, 101–114. London: CRC Press, 2018.
- PERCILLIER, Michael. „Creating and Analyzing Literary Corpora”. In *Data Analytics in Digital Humanities*, edited by Shalin HAI-JEW, 91–118.

- Multimedia Systems and Applications. Springer International Publishing, 2017. <http://www.springer.com/gp/book/9783319544984>.
- PERILLI, LORENZO e DOMENICO FIORMONTE. *La macchina nel tempo: studi di informatica umanistica in onore di Tito Orlandi*. Firenze: Le lettere, 2011.
- PÉTER, RÓBERT, ZSOLT SZÁNTÓ, JÓZSEF SERES, VILMOS BILICKI and GÁBOR BEREND. „AVOBMAT: A Digital Toolkit for Analysing and Visualizing Bibliographic Metadata and Texts”. In *XVI. Magyar Számítógépes Nyelvészeti Konferencia*, edited by Veronika VINCZE, Gábor BEREND and Gábor GOSZTOLYA, 43–55. Szeged: Szegedi Tudományegyetem, Informatikai Intézet, 2020. http://www.inf.u-szeged.hu/~berendg/docs/publ/AVOBMAT_final.pdf.
- PIERAZZO, Elena. *Digital Scholarly Editing: Theories, Models and Methods*. London–New York: Routledge, 2016. <https://doi.org/10.4324/9781315577227>.
- PLECHÁČ, Petr, Barry P SCHERR, Tatyana SKULACHEVA, Helena BERMÚDEZ-SABEL and Robert KOLÁR, eds. *Quantitative Approaches to Versification*. Praha: Institute of Czech Literature of the Czech Academy of Sciences, 2019. <http://versologie.cz/conference2019/proceedings.php>.
- ROBINSON, Peter. „Towards a Theory of Digital Editions”. *The Journal of the European Society for Textual Scholarship*, 10, Variants (2013): 105–131. https://doi.org/10.1163/9789401209021_009.
- ROCKWELL, G. and S. SINCLAIR. *Hermeneutica: Computer-Assisted Interpretation in the Humanities*. Cambridge: The MIT Press, 2016.
- ROSSELLI DEL TURCO, Roberto. „Il Vercelli Book Digitale: teoria e pratica di un progetto di edizione elettronica”. *Schede Umanistiche* 28 (2014): 181–200.
- SAHLE, Patrick. *Digitale Editionsformen: zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*. Norderstedt: Books on Demand, 2013.
- SCHMIDT, Desmond Allan. „Using Standoff Properties for Marking-up Historical Documents in the Humanities”. *Information Technology* 58, no. 2. (2016): 63–69. <https://doi.org/10.1515/itit-2015-0030>.
- SCHREIBMAN, Susan, Ray SIEMENS and John UNSWORTH, eds. *A New Companion to Digital Humanities*. 2nd edition. Malden: Wiley-Blackwell, 2016. <https://www.wiley.com/en-us/A+New+Companion+to+Digital+Humanities%2C+2nd+Edition-p-9781118680599>.
- STELLA, Francesco. „Il problema della codifica nelle edizioni filologiche digitali”. In *Nel segno del testo: Edizioni, materiali e studi per Oronzo Pecere*, a cura di Lucio DEL CORSO, Franco DE VIVO e Antonio STRAMAGLIA, 347–357. *Papyrologica Florentina* 44. Firenze: Edizioni Gonnelli, 2015.
- TOMMASI, Francesca. „L’edizione digitale e la rappresentazione della conoscenza: Un esempio: Vespasiano da Bisticci e le sue lettere”. *Ecdotica* 9 (2012): 264–286.

- UNDERWOOD, Ted. *Distant Horizons Digital Evidence and Literary Change*. Chicago: University of Chicago Press, 2019.
- WEITIN, Thomas. „Digitale Literaturwissenschaft”. *Deutsche Vierteljahrschrift für Literaturwissenschaft und Geistesgeschichte* 89, no. 4. (2015): 651–656. <https://doi.org/10.1007/BF03396502>.
- „Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden”. *Zeitschrift für digitale Geisteswissenschaften* 3 (2018). <http://www.zfdg.de/sonderband/3>.
- WILKINSON, Mark D., Michel DUMONTIER, IJsbrand Jan AALBERSBERG, Gabrielle APPLETON, Myles AXTON, Arie BAAK, Niklas BLOMBERG, et al. „The FAIR Guiding Principles for Scientific Data Management and Stewardship”. *Scientific Data* 3 (15 March 2016): 160018. <https://doi.org/10.1038/sdata.2016.18>.

HAZAI SZAKIRODALOM

- „Alapelvek az irodalmi szövegek tudományos kiadásához”. *Irodalomtörténet* 35, 3. sz. (2004): 328–330.
- ALMÁSI Zsolt. „Szöveg az olvasónak: kvantitatív módszerek és digitális Shakespeare-filológia”. *Magyar Tudomány* 177, 11. sz. (2016): 1286–1291.
- BALÁZS-HAJDU Péter, BOGNÁR Péter, HEVESI Andrea és SINKA Zsófia. „Szegedi kísérlet a 17. századi magyar vers gépi feldolgozására”. In *Filológia és textológia a régi magyar irodalomban: tudományos konferencia, Miskolc, 2011. május 25–28.*, szerkesztette KECSKEMÉTI Gábor és TASI Réka, 461–470. Miskolc: ME BTK Magyar Nyelv- és Irodalomtudományi Intézet, 2012.
- BÍRÓ Szabolcs. *Szövegfeldolgozás XML alapokon*. Neumann Kht., 2005. <http://www.tankonyvtar.hu/hu/tartalom/tkt/szovegfeldolgozas-xml/index.html>.
- BODROGI Ferenc Máté. „Elektronikus kritikai kiadások a klasszikus magyar irodalomban”. *Magyar Tudomány* 177, 11. sz. (2016): 1292–1296.
- DEBRECZENI Attila. „Kritikai kiadás papíron és képernyőn”. In *Textológia, filológia, értelmezés: Klasszikus magyar irodalom*, szerkesztette CZIFRA Mariann és SZILÁGYI Márton, 26–42. Bibliotheca Studiorum Litterarium 55. Debrecen: Debreceni Egyetemi Kiadó, 2014.
- DOBÍ JAN Sándor, MÉSZÁROS Tamás és KISS Margit. „Shtylo: stilometriai elemzések webes támogatása”. In *XIV. Magyar Számítógépes Nyelvészeti Konferencia*, szerkesztette VINCZE Veronika, 423–436. Szeged: Szegedi Tudományegyetem, Informatikai Intézet, 2018.
- FORGÁCS Tamás. „Zárt korpuszok és pótkompetenciák”. *Néprajz és Nyelvtudomány* 35, (1994): 17–23.

- FÜREDI Mihály. „A mai magyar széppróza statisztikai vizsgálata”. In *A magyar nyelv rétegződése*, szerkesztette Kiss Jenő és Szűts László, 1:336–349. 2 köt. Budapest: Akadémiai Kiadó, 1983.
- GÁL György. „A »Répertoire de la Poésie Hongroise Ancienne« adatmodellje”. *Irodalomtörténeti Közlemények* 93, 3. sz. (1989): 267–272.
- GOLDEN Dániel. „Szöveg((l)elés)”. *Palimpszeszt*, 10. sz. (1998). http://magyar-irodalom.elte.hu/palimpszeszt/10_szam/17.htm.
- GOLDEN Dániel. „Wörösmarty és a (digitális) barbárság kora”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. március 4. magyar-irodalom.elte.hu/biop/barbar/cikkek/gd.htm.
- HORVÁTH Iván. „Digitális bölcsészet a virtuális nemzeti könyvtárban”. *Digitális Bölcsészet* 1 (2018): 121–138. <https://doi.org/10.31400/dh-hun.2018.1.234>.
- HORVÁTH Iván. „Egy műfaj halála”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. március 4. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/hi.htm>.
- HORVÁTH Iván. *Gépeskönyv*. Budapest: Balassi Kiadó, 2006.
- HORVÁTH Iván. „Számítógép és irodalomtudomány”. *Alföld* 61, 3. sz. (2010): 9–28.
- JOCKERS, Matthew L. „Metaadat”. Fordította LABÁDI Gergely. *Digitális Bölcsészet* 1 (2018): 83–108. <https://doi.org/10.31400/dh-hun.2018.1.242>.
- KALCSÓ Gyula. „Digitális források – digitális filológia”. In *Módszertani tanulmányok*, szerkesztette BALLABÁS Dániel, 35–43. Eger: EKF Líceum Kiadó, 2013.
- KECSKEMÉTI GÁBOR. „A textológiai munka egyes problémáiról: az új textológiai alapelvek közrebocsátásakor”. *Irodalomtörténet* 35, 3. sz. (2004): 317–327.
- KIRÁLY Péter. „Kritikai kiadás és XML”. In *Septuagesimo anno humanissime peracto: Tanulmányok Kulcsár Péter 70. születésnapjára*, szerkesztette HORVÁTH Zita, 119–149. Publicationes Universitatis Miskolcensis: Sectio Philosophica, IX/4. Miskolc: Miskolci Egyetem, 2004.
- KISS Katalin. „Variánskiadások előkészítése TEI/XML-alapokon (Szabó Lőrinc: *Föld, Erdő, Isten*)”. In „*Szabad ötletek...*”: Szőke György tiszteletére barátaitól és tanítványaitól, szerkesztette KABDEB Lóránt, RUTTKAY Helga és SZABÓNÉ HUSZÁRIK Mária, 270–280. Miskolc: Miskolci Egyetem BTK Irodalomtudományi Doktori Iskola–Szabó Lőrinc Kutatóhely, 2005.
- KISS Margit. „Egy XVIII. századi írói korpusz modern feldolgozása”. *Magyar Tudomány* 177, 11. sz. (2016): 1305–1309.

- KISS Margit. „Stilometriai elemzés lehetőségei magyar történeti szövegkorpuszon”. *Digitális Bölcsészet* 2 (2019): T:15–T:33. <https://doi.org/10.31400/dh-hun.2019.2.336>.
- KOKAS Károly. „Az elektronikus könyvtáros textológiai gondjai”. In *Networkshop '97*. Keszthely: NIIF Koordinációs Iroda, 1997. <http://publicatio.bibl.u-szeged.hu/11523/>.
- KÓMÁR Éva, CSÉVE Anna és FELLEGI Zsófia. „Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadása”. *Digitális Bölcsészet* 1 (2018): 159–174. <https://doi.org/10.31400/dh-hun.2018.1.227>.
- KOZÁK Dániel. „Digitális szövegkorpuszok a klasszika-filológiában”. In *A humán tudományok és a gépi intelligencia*, szerkesztette TOLCSVAI NAGY Gábor, 62–91. Budapest: Gondolat Kiadó, 2018. <http://real.mtak.hu/88669/>.
- LABÁDI Gergely. „A filológiai tudás formái”. In *Textológia, filológia, értelmezés: Klasszikus magyar irodalom*, szerkesztette CZIFRA Mariann és SZILÁGYI Márton, 173–190. Bibliotheca Studiorum Litterarium 55. Debrecen: Debreceni Egyetemi Kiadó, 2014.
- LABÁDI Gergely. „Az olvasógép”. *Digitális Bölcsészet* 1 (2018): 17–34. <https://doi.org/10.31400/dh-hun.2018.1.126>.
- LABÁDI Gergely. „Digitális filológia”. In *Bevezetés az irodalomtudományba*. Szeged: Szegedi Tudományegyetem, 2013. http://www.jgypk.hu/mentorhalo/tananyag/Bevezets_az_irodalomtudomnybaV2/ii_3_digitalis_filologia_labdi_gergely.html.
- LABÁDI Gergely. „Géppel mért irodalom: a mikszáthi előbeszédszerűség”. *Digitális Bölcsészet* 2 (2019): M:3–M:19. <https://doi.org/10.31400/dh-hun.2019.2.390>.
- LABÁDI Gergely. „Könyvek távolról: A magyar regény 1807-ben”. *Irodalomtörténet* 95, 3. sz. (2014): 311–332.
- LEJTOVICZ Katalin, SCHLÖGL Matthias, BERNÁD Ágoston Zénó, KAISER Maximilian és RUMPOLT Peter Alexander. „Digitalizáció és hálózatkutatás: Az *Österreichisches Biographisches Lexikon 1815–1950* és az APIS-projekt”. *Digitális Bölcsészet* 1 (2018): 139–158. <https://doi.org/10.31400/dh-hun.2018.1.232>.
- MARKÓ Anita. „Hálózatok a 16–17. századi album amicorumokban”. *Digitális Bölcsészet* 1 (2018): 55–82. <https://doi.org/10.31400/dh-hun.2018.1.152>.
- MARÓTHY Szilvia. „A tudományos elektronikus szövegkiadások feldolgozása”. In *Valóságos könyvtár – könyvtári valóság*, szerkesztette KISZL Péter és CSÍK Tibor, 351–356. Budapest: ELTE BTK Könyvtár- és Információtudományi Intézet, 2018.

- MARÓTHY Szilvia. „Tudományos szövegkiadások a hálózaton: Áttekintés”. *Irodalomtörténeti Közlemények* 122, 5. sz. (2018): 617–633. <http://hdl.handle.net/10831/43664>.
- MESTERHÁZI Gábor. „Kultúra közpénzen?”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/mg.htm>.
- MÉSZÁROS Ágnes. „Genetikus elv – digitális kiadás (Elméleti megfontolások egy készülő Kazinczy-kiadás kapcsán)”. *Irodalomtörténeti Közlemények* 105, 3–4. sz. (2001): 414–430.
- MÉSZÁROS Tamás. „Mit nyújthat a modern informatika az irodalomtudomány számára?”. *Magyar Tudomány* 177, 11. sz. (2016): 1310–1315.
- ORLOVSZKY Géza. „A digitális szövegkiadás helyzetéről: vitaindító és javaslatok”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/og.htm>.
- PALKÓ Gábor. „Digitális filológia: számítógép anyaszerepben”. *Filológiai Közöny* 61, 2. sz. (2016): 187–199.
- PALKÓ Gábor. „Mit jelent a digitális filológia a szemantikus web korában?”. *Magyar Tudomány* 177, 11. sz. (2016): 1316–1322.
- PARÁDI Andrea. „Internetes kritikai kiadás”. In *A magyar irodalom filológiája*. Digitális Tankönyvtár–Gépeskönyv, 2005. <http://www.tankonyvtar.hu/hu/tartalom/tkt/magyar-irodalom/ch13.html>.
- PÉTER Róbert. „A Big Data kihívás és lehetőség a bölcsészettudományokban: digitális szövegek és metaadatok távoli olvasása”. *Magyar Tudomány* 177, 11. sz. (2016): 1323–1330.
- POMÁZI Gyöngyi. „Elektronikus szövegkiadás szakmai háttér nélkül?”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/pgy.htm>.
- RPHA MUNKACSOPORT. „Szegedi kísérlet a XVI. századi magyar vers gépi feldolgozására”. *Irodalomtörténeti Közlemények* 84, 5–6. sz. (1980): 630–638.
- SELÁF Levente. „Digitális versrepertóriumok fejlesztése és összekapcsolása: kutatástörténet és kilátások”. *Magyar Tudomány* 177, 11. sz. (2016): 1331–1336.
- TÓTH Tünde. „Irodalomtörténészek a bábeli könyvtárból. A régi magyar vers repertórium”. *Irodalomtörténeti Közlemények* 101, 1–2. sz. (1997): 133–145.
- TÓTH Tünde. „Bábeli nyelvzavar: Mi az a »corpus digitale«?” [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. március 4. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/tt.htm>.

- TUBAY Tiziano. „Irodalomtudomány és természettudomány”. In */HI70/ tanítóanyagok: Tanulmányok Horváth Iván 70. születésnapjára*, szerkesztette BARTÓK Zsófia Ágnes, BOGNÁR Péter és MARÓTHY Szilvia. Budapest: Q. E. D. Kiadó, 2018. <http://hi70.hu/2018/03/27/tubay/>.
- TURI László, TÓTH Tünde és GOLDEN Dániel. „Virtuális örökkévalóság: objektumok a digitális könyvtárban”. *Palimpszeszt*, 10 (1998). <http://magyar-irodalom.elte.hu/vita/>.
- VADAI István. „Szövegrögzítés”. [Előadás.] *Filológia és digitális barbárság*, Budapest: ELTE BTK Bölcsészettudományi Informatika Önálló Program, 2004. március 4. <http://magyar-irodalom.elte.hu/biop/barbar/cikkek/vi.htm>.
- ZSOLDOS-DEMJÉN Anetta. „Digitális szövegrögzítés TEI irányelvek alapján”. In *Filológia és textológia a régi magyar irodalomban: tudományos konferencia, Miskolc, 2011. május 25–28.*, szerkesztette KECSKEMÉTI Gábor és TASI Réka, 491–504. Miskolc: ME BTK Magyar Nyelv- és Irodalomtudományi Intézet, 2012.

HELIKON TEMATIKUS SZÁMOK

- KAPPANYOS András, szerk. „A hipertext”. *Helikon* 50, 3. sz. (2004).
- NÉMETH Zoltán és VÁSÁRI Melinda, szerk. „Hálózatelmélet és irodalomtudomány”. *Helikon* 63, 2. sz. (2017). https://iti.btk.mta.hu/images/kiadvanyok/helikon/Helikon_2017_2.pdf.
- PALKÓ Gábor, szerk. „Az archívumok elméletei”. *Helikon* 60, 3. sz. (2014).
- PETŐFI S. János, szerk. „A számítógépek és a humán tudományok (vegyes szám)”. *Helikon* 15, 3–4. sz. (1969).

Összeállította: Maróthy Szilvia

KITEKINTÉS

TÜSKÉS GÁBOR

Télemakhosz német nyelvterületen a 18. században

Christoph SCHMITT-MAASS. *Fénelons „Télémaque” in der deutschsprachigen Aufklärung (1700–1832)*. 2 Bände. Frühe Neuzeit 220. Berlin–Boston: De Gruyter, Edition Niemeyer, 2018. 1289, 6 ill.

Fénelon *Télémaque*-ja a 18. század egyik sikerkönyve, melynek elterjedése az európai irodalom és művelődés megkerülhetetlen fejezete. A több művészeti ágat érintő elterjedést, melynek kezdeteit Pierre Valentin Faydit már 1700-ban a *Télémacomanie* kifejezéssel illette, számos tényező ösztönözte. Érdemes közülük kiemelni a különböző műfaji minták mesteri összekapcsolását, a lelki folyamatok és érzelmek árnyalt ábrázolását, az előremutató politikai és más eszmék markáns jelenlétét, a természet idillikus szemléletét, az erkölcsi tanítás tetszetős formába öltöztetését, a választékos, elegáns és könnyed elbeszélésmódot, a kulcsregény-jelleget, a szerző kegyvesztettségét és végül a mű betiltását. A robbanásszerű elterjedés és rendkívüli népszerűség összetevői nyelvterületenként részben eltértek egymástól, s ez is hozzájárult ahhoz, hogy az utóbbi évtizedekben megszorodtak a regény recepciójával kapcsolatos, egy-egy nyelvterületre összpontosító vizsgálatok.

A korai német kutatások többnyire szűk forrásbázisra épültek, figyelmen kívül hagyták a fordítások és adaptációk mint önálló esztétikai teljesítmények és kulturális közvetítők vizsgálatát, s nem foglalkoztak a francia nyelvű kiadások német nyelvterületen végbement elterjedésének kérdéseivel. Ezen a helyzeten próbáltak változtatni a 20. század utolsó harmadától Volker Kapp, a német *Télémaque*-kutatás doyenjének a mű forrásait feltáró és a történeti kontextualizálás fontosságát hangsúlyozó vizsgálatai, kutatástörténeti áttekintései. 2001-ben megjelent monográfiájában Olaf Simons a recepció korai, 1710 és 1720 közötti szakaszára összpontosítva könyv- és kultúrtörténeti kérdéseket tett fel, s a befogadói motivációkat és érdekeltségeket állította a középpontba.

Christoph Schmitt-Maaß kétkötetes monográfiájában kettős célt tűzött maga elé: 1) módszeresen számba venni a 18. századi német recepció formáit, ösztönzőit, eszmei kontextusait; 2) összefüggően bemutatni a folyamat időbeli alakulását és a társadalmi hordozó rétegek szerepét. Alapkérdései, melyekre választ keresett, csak látszólag egyszerűek: 1) kik és hogyan olvasták a *Télémaque*-ot a 18. században a német nyelvterületen? 2) hozzájárult-e a mű recepciója a felvilágosodás, az

új politikai, történeti diskurzusok és az irodalmi ízlésváltozás kibontakozásához?
3) ha igen, hogyan és mennyiben, s milyen szerepet játszott e folyamatokban?

A monográfia elsődleges forrásanyagát a német nyelvterületen 1700 és 1806 között önállóan megjelent francia és kétnyelvű *Télémaque*-kiadások, az ugyanitt közreadott német, olasz, angol és latin nyelvű fordítások, adaptációk, részfordítások, kivonatok, drámaprogramok, librettók és szótárak teljességre törekvő, közel százhatvan tételes katalógusa adja. A felhasznált kéziratokat és az elemzésbe bevont további nyomtatott forrásokat, melyek összesen mintegy kilencven (!) könyvtárban és levéltárban található, egy közel hatvan lap terjedelmű jegyzék tartalmazza. A kéziratok között találunk például kölcsönzési naplókat, könyvjegyzékeket, katalógusokat, levelezéseket, jezsuita *litterae annuae*-ket, tanterveket, nevelési programokat és instrukciókat. A nyomtatott források összetétele hasonlóképpen változatos. Figyelemre méltó, hogy a két kötet összterjedelmének kb. a hatodrésztét a katalógus, a forrás- és irodalomjegyzék, valamint a mutatók teszik ki. A forrásanyag feldolgozását harminckilenc táblázat segíti, melyek többsége részletes szövegösszevetés alapjául szolgál. A vizsgálat térbeli kereteit a bevezető szerint a Német-római Birodalom 1790 előtti határai jelölték ki, így Poroszország, Holstein, Szilézia és Csehország mellett beletartozik a Habsburg Birodalom, beleértve Tirolt, továbbá a németalföldi Habsburg-tartományok, a svájci kantonok államszövetsége és Baden, beleértve a Pfalzi Választófejedelemséget.

A recepció történetét Schmitt-Maaß három nagy részre – 1700 előtt; 1700–1790 között; 1790–1832 között –, ezen belül a második részt ugyancsak három, harmincéves szakaszra, s az egész munkát a bevezetővel együtt összesen tizenöt fejezetre osztva mutatja be. A második rész harmincéves szakaszokra bontását azzal indokolja, hogy a szakaszok egy-egy generációnyi időt fognak át, és igazodnak a német *Télémaque*-fordítások megjelenéséhez. Az első szakaszt (1730-ig) – némileg mechanikusan – a felvilágosodásnak a korai, a másodikat (1760-ig) a középső, a harmadikat (1790-ig) a késői periódusával kapcsolja össze. A részek tagolását erősen befolyásolta a társadalomtörténeti szemlélet: a III. rész kivételével a szerző minden szakaszon belül külön tárgyalja a fejedelmi udvarok, az egyházi körök, a tudósok és a polgári rétegek recepciójának megnyilvánulásait. A III. részben hat fejezetben mutatja be a weimari udvarhoz kapcsolódó 18. századi dokumentumokat, két hercegi könyvtár, a polgári olvasókörök és az alsó rétegek olvasmányait tükröző kölcsönkönyvtárak tanúságát, a Bécs–Weimar viszonylatban 1797–1805 között lezajlott opera-vitát, a weimari klasszika meghatározó alakjainak *Télémaque*-képét, egy női olvasó ironikus átértelmezését és Fénelon 1800 körüli újrafelfedezését. Ez a szerkezet jól áttekinthető, ugyanakkor némileg megnehezíti a különböző rétegek és csoportok recepciója közti „áthallások” és közvetítési folyamatok nyomon követését.

A módszer sajátossága, hogy a germanisztikai látószöveget Schmitt-Maaß következetesen kitágítja a francia forrásszövegek és a fordítások kapcsolata felé, s figyelmét kiterjeszti a korábban másodlagos jelentőségűnek ítélt fordítókra, a

receptió okaira, eszközeire, módozataira és eszmei összefüggéseire. A hagyományos komparatista befogadástörténetet könyv-, könyvtár- és olvasástörténeti, társadalom- és kommunikációtörténeti, valamint esztétikai szempontokkal kombinálja, s eredményesen alkalmazza a kulturális transzfer és kulturális fordítás antropológiai koncepcióit. Bemutatja a fordítók szellemi profilját, munkásságát és más fordításait, elemzi a fordítások paratextusait, s szövegösszevetések révén feltárja a fordítói módszereket, nyelv- és irodalomszemléleti, poétikai sajátosságokat. Megvizsgálja a receptió résztvevőinek státuszát, szándékait és érdekeltségét, elemzi az elsajátítás formáit, korlátait és következményeit.

A könyv fontos eredménye, hogy minden eddiginél pontosabban kirajzolódnak a 18. századi német *Télémaque*-receptió sajátosságai. Fénelon már 1700 előtt is jelen volt a nyelvterületen, többek között Elisabeth Charlotte d'Orléans pfalzi rokonságával folytatott levelezésében és pietista körökben; Leibniz külön értekezésben foglalkozott Fénelon *amour désintéressée*-koncepciójával. A fejedelmi udvarok hosszú időn át a szelektív befogadás gyakorlatát követték: a berlini udvarban 1700 körül készült kéziratos kivonat és nyomtatott fiktív dialógus például határozott pedagógiai programba illeszkedik: mindkettő mellőzi a cselekményt, nem veszi tudomásul a regény abszolutizmus-kritikáját és egy új államberendezkedésre vonatkozó javaslatait, s figyelmen kívül hagyja Fénelon szándékát a klasszikus műveltségi javak és a quietista elképzelések narratív közvetítésére. Míg August Bohse 1700-ban megjelent, az udvari és polgári közönségnek egyaránt szánt fordítása a gáláns regény szférájába ültette át Fénelon stíluseszményét, Benjamin Neukirch verses fordítása (1727/39) a barokk udvari-heroikus regény mintáját követi, s a verses eposz korai német példájaként értékelhető. A fejedelmi udvarokban 1706–1722 között bemutatott operák többnyire ugyancsak háttérbe szorítják az abszolutizmus reformját célzó tartalmakat, s az *honnête homme* gáláns ideálját népszerűsítik. A tudós körök reflektálnak ugyan a mű poétikai státuszáról és műfaji hovatartozásáról folytatott francia vitára, de a vita itt szűkebb mederben folyik, ugyanakkor a reflexiók közvetve előkészítik a *Télémaque* polgári értelmezését. Carl Arnd, Leibniz és Ehrenfried Walther v. Tschirnhaus abszolutizmusellenes állásfoglalásként és a különböző államformák megismerését szolgáló új pedagógiai modell előképeként értékeli a regényt, mások természetjogi relevanciáját hangsúlyozzák. A polgári rétegeknél 1730 előtt a *Télémaque* a gáláns viselkedés kézikönyve, az iskolai és egyetemi nyelvoktatás eszköze, s megjelenik több erkölcsi hetilap női olvasóknak szóló olvasmányajánlatában is.

Az elemzés másik tanulsága, hogy az 1730 és 1760 közti évtizedekben felerősödnek a korábbi receptió formái, egyben végbemegy azok revíziója. II. Frigyes *Antimachiavell* (1740) című írása például már egy megváltozott uralkodófelfogást tükröz, melynek egyik forrása a *Télémaque*. Érvényesül a mű hatása a kertművészetben és a képzőművészetben, több latin fordítás készül, s megszorodnak a színpadi adaptációk. A receptió aktív hordozói között megjelenik a pietista klérus és a jezsuita rend; Franz Neumayr 1751-ben kiadott poétikája kifejezetten ajánlja a

regény színpadi feldolgozását. Megnő a tudós körök szerepe: a Lipcse és Zürich között zajló irodalmi vitában, mely az 1740-es években több részfordítást ösztönzött, a mű erkölcsi tartalmai mellett szerepet kapnak a formai problémák. Hős-eposz-poétikájában Gottsched az egyik fő hivatkozási ponttá emeli Neukirch fordítását. Új vonás, hogy a regény műfajjal összefüggésben is tárgyalják a művet, s az eredeti francia szöveghez nyúlnak vissza az érvelésben. A polgári körökben a jegyzetelt kiadások és fordítások révén egyre meghatározóbbá válik a nyelvtanulás motivációja. Lessing színházelméletében és részvétfogalmában – Diderot közvetítésével – a *Télémaque* Philoktetosz-epizódja kap szerepet.

Schmitt-Maaß fontos megfigyelése, hogy az 1760–1790 közti évtizedekben a fejedelmi udvarok körében előtérbe lép a szórakoztató funkció, elsősorban balett-adaptációk és operafeldolgozások formájában. A Habsburg Birodalomban felerősödik a katolikus recepció a pedagógia megújítását célzó törekvések keretében: az északnémet tankönyv-kiadások utánnomásai mellett önálló tankönyvek készülnek, s a *Pietas*-drámákban a világi szerelemtől való óvás gondolatán kívül feltűnik a hazaszeretet motívuma. Albrecht v. Haller *Usong* (1771) című, konzervatív szemléletű államregénye a *Télémaque* politikai erényprogramját aktualizálja, egyben megnyitja a műfajt az utópikus narratívák felé. Új fordítások készülnek, részben nevelési célzattal, míg a melodráma-adaptációk a lélektani-érzelmi összetevőket hangsúlyozzák.

A weimari klasszika és romantika képviselőinek érdeklődését tanúsítja Schmitt-Maaß szerint, hogy Wieland *Der Goldne Spiegel* (1772) című nevelési regénye lényegében a *Télémaque* pedagógiai programját kommentálja és bírálja. Goethe hivatkozása Neukirch fordítására Klopstock *Messiása* méltatásának keretébe illeszkedik. Herder költői szempontból elismerte, politikai vonatkozásban azonban kevésbé értékelte a *Télémaque*-ot, miközben elkészítette annak mindeddig nem közölt részfordítását. Schiller 1786-ban travesztiát írt a műre, ugyanakkor a *Don Carlos* és a *Filozófiai levelek* szeretet- és barátságfelfogása, boldogságkonceptiója egyaránt Fénelonhoz kapcsolódik. A „misztikus” Fénelon 1800 körül végbement újrafelfedezése keretében a *Télémaque*-receptiót fokozatosan felváltja a teológus Fénelon értékelése, akit igyekeznek különféle missziós szándékok, ébredési mozgalmak és irenikus törekvések szolgálatába állítani.

Összegezve megállapítható, a monográfia átfogó és igen részletgazdag képet rajzol a német nyelvterület *Télémaque*-receptiójáról, s meggyőzően tanúsítja annak sokrétűségét, megújulási képességét. Jelzi a különböző befogadói csoportok szerepének és érdeklődésének változását, szemlélteti a magasabb rétegek mintáinak átvételét az alacsonyabb társadalmi körökben, s érzékelteti a receptió súlypontjainak a század közepétől végbement elmozdulását a fejedelmi udvaroktól a polgárság felé. Megfigyelhető az is, hogy a regény receptiója közvetve ösztönözte Fénelon más, teológiai és pedagógiai tárgyú írásainak befogadását. Bár Schmitt-Maaß nem ismeri Köpeczi Béla *Télémaque*-tanulmányait, más szakirodalom alapján a katalógusban feltünteteti Gregor Trautwein latin fordításának két

kassai kiadását, jegyzetben hozza Haller László és Zoltán József magyar fordításainak kiadásait, tud a jezsuita Johann Zimmermann 1750–51-ben Kassán megjelent önálló latin fordításáról, s néhány további magyar vonatkozású adatot is közöl. Csak remélhető, hogy a monográfia mielőbb ösztönözni fogja a magyarországi *Télémaque*-repció ehhez hasonlóan módszeres feltárását és beillesztését a nemzetközi összefüggésekbe.

KÖNYVEK

Véronique Le Ru. *Émilie du Châtelet philosophe.* Les Anciens et les Modernes – Études de philosophie 38. Paris: Classiques Garnier, 2019. 206.

Véronique Le Ru könyvének célkitűzése, hogy Émilie du Châtelet-t mint filozófust mutassa be. A szerző először tömören összefoglalja azokat a vádakot és kritikákat, amelyek M^m du Châtelet-t megfosztották a tudós és filozófus státuszától, akit – mint hangsúlyozza – nem csupán Newton fordítójaként kell elismernünk, s főképp nem Voltaire szeretőjeként emlékeznünk rá, hanem megérdemli, hogy önmagáért olvassuk. A bevezetésben röviden ismerteti a márkiné családi háttérét, korai tanulmányait, kapcsolatát Voltaire-rel, azonban kiemeli, hogy M^m du Châtelet fiatalon önálló tudományos munkába kezdett.

A könyv két részből áll, melyek közül az első Émilie du Châtelet életművét vizsgálja, a második életművének fogadtatását. Az első rész első fejezete M^m du Châtelet részvételét értékeli a mozgó testekre ható erő körüli vitában (*vis viva vita*), elsősorban állásfoglalásának metafizikai hátterére koncentrálnak. A következő fejezet az *Institutions de physique* [*A fizika alapjai*, 1740] című munkát tárgyalja: Le Ru kiemeli, hogy a newtoni fizika és leibnizi metafizika összeegyeztetési kísérlete igen eredeti vállalkozás volt. Ez egy viszonylag ismert nézet, Le Ru azonban jól követhető magyarázatokkal bontja ki, hogy miben állt ez a kísérlet: magyarázatot adhatnak-e a newtoni tömegvonzás elvre, amelyet maga Newton jelenlegként vizsgált, Leibniz metafizikai elvei? Le Ru elemzi az *Institutions* Előljáró beszédében kifejtett tudományos módszert, az autoritás felülvizsgálását és a hipotézisek (akár téves hipotézisek) jótékony szerepét. Egy rövidebb fejezet a szellem éleslátásának (*sagacité*) és rugalmasságának (*flexibilité*) fogalmait vizsgálja Émilie du Châtelet gondolkodásában, melyeket a márkiné a tudományos autoritás és pártoskodás ellenében méltat. A következő fejezetben Le Ru *A boldogságról* című esszének szenteli figyelmét: a szenvedélyek elfogadása, a sztoikus halálfilozófia elvetése és egyfaj-

ta materialista (nem üdvözülésközpontú) boldogságkép meghatározó tényezőik. A boldogság nemi szerepek szerinti megkülönböztetése (*conception genrée du bonheur*) is fontossá válik, a nők számára Émilie du Châtelet a csábítás, a tetszeni vágyás helyett az önképzés programját hirdeti meg (*s'instruire au lieu de séduire*). A szerző külön fejezetet szentel az *Examens de la Bible*-nek: B. Schwarzbach kutatásai nyomán kiemeli, hogy Émilie du Châtelet a nagyívű bibliakritikai munka önálló szerzője, amely az *Institutions* első megjelenése után íródott. A munka legmarkánsabb vonásai a vallási autoritással, illetve a vallási alapú nőgyűlölettel való szembehelyezkedés. Hosszú fejezet foglalkozik Newton *Principiájának* fordításával (posztumusz, 1756 és 1759) és a Newton-mű harmadik könyvének önálló kommentárjával. Le Ru más szakértők kutatásait folytatva hangsúlyozza, hogy a fordítás nem csupán a latin átültetése franciára, hanem a szöveg átvitele az analitikus matematikai nyelvezetre, melynek köszönhetően a fordítás napjainkban is használható.

A könyv második részének első fejezete az *Institutions* németországi fogadtatását vizsgálja, valamint azt, hogy az enciklopédisták miként mértették belőle. Émilie du Châtelet munkáit a lipcsei *Gesellschaft der Aletophilien* [Igazságszeretők társasága] és a Gottsched házaspár méltatta, elsősorban Wolff filozófiájának francia hívét látták a márkinében. Samuel Formey ebben a társaságban figyelt fel a márkiné munkáira, melyek az ő közvetítése révén hatottak az *Enciklopédia* jelentős tudományos és ismeretelméleti szócikkeire. Le Ru azonban egy másik hatást is bizonyít: míg Formey szövegszerűen vesz át az *Institutions* egyes fejezeteiből, D'Alembert módszertani, tudományfilozófiai észrevételeiben mutatható ki hatása. Egy igen érdekes fejezet foglalkozik a márkiné gondolatainak rejtett fogadtatásával (*postérité souterraine*). Későbbi tudósok közül Henri Poincaré és Claude Bernard módszertani és tudományfilozófiai fejtegetéseiben lehet kimutatni a hatását, ők maguk azonban valószínűleg nem tudták, hogy az *Enciklopédia* egyes szócikkei, köztük a hipotéziselmélet,

M^{me} du Châtelet-től merítették. Ennél még érdekesebb Gaston Bachelard: ő jól ismerte a márkiné írásait, és programja – a tudomány az ellentmondások kiküszöbölése révén, megcáfolt vagy igazolt hipotézisek által fejlődik – több ponton követi az *Institutions* szerzőjét. Bachelard mégis leszólja a márkinét, és a tudomány előtti szellemnek (*esprit préscientifique*) minősíti. Külön fejezet foglalkozik La Mettrie materialista-utilitarista boldogságkonceptiójának és Émilie du Châtelet boldogságírásának egymásra tett hatásával. A márkiné több ízben eszmét cserélt La Mettrie-vel, ám újszerűsége vele szemben, hogy a nő boldogságát célzottan vizsgálja, nézeti azonban kevésbé utilitaristák. Az utolsó előtti fejezet M^{me} du Châteletnek a nők neveléséről alkotott gondolatainak utóéletét vizsgálja a forradalmi terror idején lefejezett Olympe de Gouges-nál. Nem tudjuk, hogy e szerző honnan és milyen mélységig ismerte a márkiné írásait. Gondolatmenete az alábbi szempontból viszi tovább elképzeléseit: Rousseau-val ellentétben azt hangsúlyozza, hogy a természet nem rendeli alá a nőt, a nő alárendeltsége társadalmi. Tanulva, magát képezve a nő megszabadul a csábítás és intrika által szerzett befolyás méltatlan szerepétől. Az utolsó fejezet az Émilie du Châtelet-vel szembeni ellenérzést választja tárgyként. Meglephet bennünket, hogy Le Ru nem a márkiné halálát követő, gyakran gúnyos sajtóhíreket tekinti át, nem is a vele kapcsolatos 19. századi kritikát, hanem a luxus kérdését elemzi Voltaire és az enciklopédisták gondolkodásában, illetve azt, hogy miért válhatott ez a gondolatmenet a konzervatív kritika tárgyává. A könyv konklúziója igen eredeti hangvételű írás Émilie du Châtelet életének és személyiségének morális nagyságáról.

Le Ru könyve igen hasznos azok számára, akik a 18. század női gondolkodóival kívánnak foglalkozni. Komoly kutatómunkán alapul, pontos, jól felépített, Émilie du Châtelet személyisége és filozófiája iránti őszinte tiszteletről árulkodik. Lendületes, de kifinomult stílusban vezeti olvasóját az életművön keresztül. Le kell azonban szögezni, hogy nem könnyen olvasható munka, bár szintézisnek szánták: az olvasónak rendelkeznie kell a kora újkori tudomány és filozófia alapfogalmaival, fő vitáinak kontextusát legalább érintőlegesen ismernie kell.

KOVÁCS ESZTER

Bódi Katalin. *Éva születése.* Új Alföld Könyvek 3. Debrecen: Méliusz Juhász Péter Könyvtár, 2019. 280.

A művészet- és irodalomszerető olvasóközönség régóta vár egy olyan, magyar nyelvű könyv megjelenésére, amely az irodalom- és a kultúrtörténet kapcsolódási pontjainak szemszögéből vizsgálja a képzőművészeti alkotásokat. Bódi Katalinnak az Új Alföld Könyvek sorozatában megjelent tanulmánygyűjteménye minden szempontból megfelel e várakozásnak: az igényes kiállítású kötetet előszó, tizenhárom tanulmány, névmutató és egy rendkívül gazdag, mintegy száz illusztrációt tartalmazó színes képmelléklet alkotja. Az idegen és gyakran érthetetlen terminusoktól mentes, szép magyar nyelven megírt könyv a művészetelméleti fejtegetéseken túl konkrét műalkotások elemzését tárja az olvasó elé: abba az irányzatba illeszkedik, amelyet Magyarországon elsősorban Földes Györgyi testírással (*écriture du corps, body writing*) – a test és a szöveg rendkívül összetett kölcsönviszonyával – kapcsolatos kutatásai képviselnek (itt talán érdemes megjegyeznünk, hogy a *Helikon* folyóirat 2011/1–2. tematikus különszáma összefoglaló áttekintést nyújt az ezen a téren folyó kutatások eredményeiről).

A 18. századi regényirodalommal foglalkozó Bódi Katalin munkája ezúttal nem irodalmi, hanem képzőművészeti alkotások elemzéséből indul ki. A könyv címe és a borítóján látható kép – a Barbie babát formázó, fej nélküli, hófehér női büszt, amelyen egyedül a mellék hangsúlyosak – gondolattársítások egész láncolatát indítja el. A cím láttán az olvasónak óhatatlanul Sandro Botticelli jól ismert festménye, a *Vénusz születése* jut az eszébe. Éva születése – Vénusz születése volna? Ez az áthallás korántsem véletlen: a kötetben szereplő írások témája ugyanis az újkori európai képzőművészetben megjelenő nőábrázolás.

A képelemzések a reneszánsz művészetben gyökerező testábrázolás kettős kötődését, az antik és a keresztény testfogalom egymáshoz való viszonyát fejtegetik. Bódi Katalint a női test: a szenvedélyes, szenvedő, feltárukozó és bezáruló női test képi megjelenítésének problémája érdekli, amint az a különböző művészeti ágakhoz tartozó alkotásokban megnyilvánul. E női testtel kapcsolatban a képzőművészetek, az irodalom és az

antropológia határmezsgyéjén elhelyezkedő tanulmányok olyan kérdéseket vetnek fel, amelyeket az olvasó – ha nem is tabutémáknak, mindenestre – transzgresszívnek érzékel. A könyvön azonban az említett, explicit tematikus vonal mellett más, rejtettebb szálak is végighúzódnak, olyanok, mint például a képek narrativitásának, az ábrázolt pillanat kiválasztásának vagy a figuratív festészetben a női tekintet sajátosságának a kérdése. Olvashatunk a könyvben „női mesiasokról”, a firenzei festőiskola első női tagjáról, Artemisia Gentileschiről vagy az alkimista Médeiaról: a tanulmányok kétségkívül női látásmódot tükröznek, de az elemzések hangneme mindvégig objektív marad, tudatosan távolságot tart a manapság divatos feminista irányzatoktól.

Két esszé kivételével – amelyek kortárs képzőművészeti kiállításokra reflektálnak – az írások az újkori testábrázolás tárgykörét járják körül. A könyv alaptézise a reneszánsz művészet és az európai kultúra kettős (antik és keresztény) ihletettsége, ami évszázadokon keresztül meghatározta a művek témaválasztását. Az esszék általában olyan témákkal foglalkoznak, amelyek ábrázolása problematikus. A címadó tanulmány például abból a kérdésből indul ki, hogy a bibliai témájú festészetben szinte megoldhatatlan problémát jelent Ádám és Éva teremtésének szöveghű ábrázolása, vagyis annak a jelenetnek a képi reprezentációja, amikor Isten – Ádám oldalbordáját kihatolva – megteremti Évát. De találunk a kötetben olyan esszét is, amely a képi ábrázolás felforgató formáihoz tartozó megnyűzött, felnyíló testre összpontosítva

a passióábrázolás hagyományait helyezi új megvilágításba. Az antik gyökerű témák közül ugyancsak a képi megjelenítés nehézségére hívja fel a figyelmet a kéj és a szégyen ábrázolása, ezeket Bódi Katalin Agnolo Bronzino, illetve a Lucretia-témájú festmények kapcsán idézi fel. Különösen nyugtalanító témát dolgoz fel a nagybányai festőiskola cigányaktjaival foglalkozó esszé, amely jól példázza, hogy a könyv szerzője újszerű módon közelíti meg a nem szokványos témákat. Úgyszintén meglepő témát tárgyal a *Heverészes esztétikája* című tanulmány, amelyből megtudjuk, miért adódik hozzá Madame Récamier portréihez – mintegy attribútumként – a rekamié nevű bútordarab.

A kötetben található írások a testábrázolás kutatásának új irányait vázolják fel, és szokatlan perspektívákat nyitnak meg: a felidézett képzőművészeti alkotások egymással is párbeszédet folytatnak, több – sokszor egymást keresztező – szálon futó, izgalmas hermeneutikai játékba bonyolódnak. Sajnálhatjuk, hogy a tanulmányok között néhol feltűnő ismétlődések és átfedések vannak, aminek nyilvánvalóan az az oka, hogy a könyvbe beavogatott írások eredetileg különálló tanulmányokként láttak napvilágot. Az egy kötetbe gyűjtött, nagyívű kultúrtörténeti esszék mindazonáltal egytől egyig élvezetes olvasmányok: legfőbb erényük abban fogalmazható meg, hogy nem akarnak egyértelmű válaszokat adni a felvetett kérdésekre, inkább azok továbbgondolására készítetik az olvasót.

BARTHA-KOVÁCS KATALIN

TARTALOM

Számítógépes irodalomtudomány

TANULMÁNYOK

MARÓTHY SZILVIA: A vizuális és az adatközpontú szemlélet. Bevezető gondolatok	5
ALMÁSI ZSOLT: Számítógépes irodalomtudomány: az adattól a transzcendentális kritikáig	16
GOLDEN DÁNIEL: Számítógép, szöveg, tudomány	32

INTERJÚ

<i>hic Rhodus, hic salta</i> : Tito Orlandi és Julianne Nyhan beszélgetése (Fordította: Maczelka Csaba)	53
--	----

MŰHELY

KALCSÓ GYULA: Kézírtos források tudományos elektronikus kiadása	67
FELLEGI ZSÓFIA – PALKÓ GÁBOR: Arany-kéziratok és kritikai kiadások közzététele az Arany János Emlékévből	82

SZEMLE

SZÉNÁSI ZOLTÁN: Digitális átalakítás. A tudományos szövegkiadás lehetőségei a weben	99
--	----

VÁLOGATOTT BIBLIOGRÁFIA	113
-------------------------	-----

KITEKINTÉS

TÜSKÉS GÁBOR: Télemakhosz német nyelvterületen a 18. században	123
--	-----

KÖNYVEK

VÉRONIQUE LE RU: <i>Émilie du Châtelet philosophe</i> / KOVÁCS ESZTER	129
BÓDI KATALIN: <i>Éva születése</i> / BARTHA-KOVÁCS KATALIN	130

CONTENTS

Computational literary studies

STUDIES

SZILVIA MARÓTHY: The visual and the data-centric approach. Preliminary thoughts	5
ZSOLT ALMÁSI: Computational literary studies: from data to transcendental critique	16
DÁNIEL GOLDEN: Computer, text, criticism	32

INTERVIEW

<i>hic Rhodus, hic salta</i> : Tito Orlandi and Julianne Nyhan (Translated by Csaba Maczelka)	53
---	----

WORKSHOP

GYULA KALCSÓ: Scholarly digital edition of manuscripts	67
ZSÓFIA FELLEGI – GÁBOR PALKÓ: Publishing manuscripts and printed critical editions of János Arany digitally for the János Arany Anniversary [Arany János Emlékév]	82

REVIEW

ZOLTÁN SZÉNÁSI: Digital transformation. Opportunities of scholarly editions on the web	99
--	----

SELECTED BIBLIOGRAPHY	113
-----------------------	-----

BOOKS	123
-------	-----

SOMMARIO

Filologia computazionale

SAGGI

SZILVIA MARÓTHY: Approccio visivo e incentrato sui dati. Riflessioni preliminari	5
ZSOLT ALMÁSI: Filologia computazionale: dai dati alla critica trascendentale	16
DÁNIEL GOLDEN: Computer, testo, critica	32

INTERVISTA

<i>hic Rhodus, hic salta</i> : Dialogo tra Tito Orlandi e Julianne Nyhan (tradotto in ungherese da Csaba Maczelka)	53
--	----

WORKSHOP

GYULA KALCSÓ: Edizione critica digitale di fonti manoscritte	67
ZSÓFIA FELLEGI – GÁBOR PALKÓ: Pubblicazione dei manoscritti ed edizioni critiche di Arany nell'anno commemorativo dedicato a János Arany	82

RECENSIONE

ZOLTÁN SZÉNÁSI: Digitalizzazione – Possibilità di pubblicare edizioni critiche sul web	99
--	----

BIBLIOGRAFIA SELEZIONATA	113
---------------------------------	-----

LIBRI	123
--------------	-----

HELIKON
IRODALOMTUDOMÁNYI SZEMLE

1955–1962

Vegyes tartalmú számok

1963

1. sz. A komplex összehasonlító kutatások elvi kérdései
2. sz. Nemzetközi Összehasonlító Konferencia (Budapest, 1962)
3. sz. Amerikai prózairodalom
4. sz. Viták a realizmusról

1964

1. sz. Az összehasonlító irodalomtudomány nemzetközi szemléje
- 2–3. sz. A kelet-európai avantgárd
4. sz. Shakespeare-évforduló (vegyes szám)

1965

1. sz. Mai világirodalmi mozgalmak és irányok
2. sz. A szocialista realizmus kérdéseiről
3. sz. Nacionalizmus és kozmopolitizmus; eredetiség – utánzás – hatás fogalmai (Az AILC IV. Kongresszusa – Fribourg, 1964. – előadásaiból)
4. sz. A kelet-európai összehasonlító irodalomtörténet kérdései

1966

- 1–2. sz. Irányzatok és csoportok az 1920–30-as évek szovjet irodalmában
3. sz. Eszmék és művek a modern polgári irodalomban
4. sz. Irodalom és szociológia

1967

1. sz. Irodalom és folklór
2. sz. Pártosság, elkötelezettség, elkötelezetlenség
- 3–4. sz. A szovjet irodalomtudomány legújabb eredményeiből

1968

1. sz. A strukturalizmusról
2. sz. Az irodalmi irányzatok mint nemzetközi jelenségek (Az AILC V. Kongresszusa – Belgrád, 1967. – anyagából)
- 3–4. sz. Az irodalom és a társzművészetek

1969

1. sz. Kelet-európai irodalmak a századfordulón
2. sz. Művészet – tömegkultúra – irodalom
- 3–4. sz. A számítógépek és a humán tudományok (vegyes szám)

1970

1. sz. A Fekete-Afrika irodalmáról
2. sz. Irodalom és összehasonlító módszer (vegyes szám)
- 3–4. sz. Modern stilsztika

1971

1. sz. Irodalom és társadalom (AILC VI. Kongresszus. Bordeaux, 1970)
2. sz. Irodalomelméleti viták Franciaországban
- 3–4. sz. A közép-európai humanizmus kérdései (Sopron, 1971)

1972

1. sz. Science fiction (a műfaj esztétikai és poétikai kérdései)
2. sz. Klasszikusaink és Európa
- 3–4. sz. A szocialista országok irodalmának másfél évtizede

- 1973
1. sz. Műelemzés és műfajelmélet (vegyes szám)
 - 2 – 3. sz. Irodalomtudomány és szemiotika
 4. sz. A XVIII. század és a felvilágosodás irodalma
- 1974
1. sz. Az AILC kanadai kongresszusa (Montreal, 1973. augusztus 13–19.)
 2. sz. Az elsüllyedt kultúrák irodalma
 - 3 – 4. sz. Modern poétika
- 1975
1. sz. Irodalom, világirodalom, nemzeti irodalom
 2. sz. Az újabb Délkelet-Európa-kutatások
 - 3 – 4. sz. Az európai romantika
- 1976
1. sz. Szubkultúra és underground
 - 2 – 3. sz. Irodalom és irodalomtörténet Ausztriában
 4. sz. Tudomány-e az irodalomtudomány?
- 1977
1. sz. A retorika újjászületése
 2. sz. Az AILC VIII., budapesti kongresszusa (1976. augusztus 12–17.)
Különböző kultúrákban eredő irodalmak kapcsolatai a XX. században
 3. sz. Az AILC VIII., budapesti kongresszusa.
Összehasonlító irodalomtudomány és irodalomelmélet
 4. sz. A budai Egyetemi Nyomda szerepe a kelet-európai népek társadalmi, kulturális és politikai fejlődésében (1777–1848)
Budapest, 1977. szeptember 5–8.
- 1978
- 1 – 2. sz. Kutatási irányok a 20-as évek szovjet irodalomtudományában
 3. sz. Érték és társadalom. Transzgresszív elemzések
 4. sz. Új magyar világirodalom-történet
- 1979
- 1 – 2. sz. Az ázsiai népek irodalma
 3. sz. A jugoszláv népek irodalma
 4. sz. Az egyéni és a kollektív a nyelvben és az irodalomban
(FILLM XIV. Kongresszus, Aix-en-Provence, 1978)
- 1980
- 1 – 2. sz. Recepciókutatás és befogadásesztétika
 - 3 – 4. sz. Az orosz szimbolizmus
- 1981
1. sz. Az irodalom klasszikus modelljei – Az irodalom és a társművészetek –
A regény fejlődése
 - 2 – 3. sz. Régi és új hermeneutika
 4. sz. Irodalom és felvilágosodás
- 1982
1. sz. A Vormärz-irodalom és néhány magyar vonatkozása
 - 2 – 3. sz. Új kutatási irányok a szovjet irodalomtudományban
 4. sz. Művelődéstörténet és Kelet-Európa
- 1983
1. sz. Az AILC X. Kongresszusa
 2. sz. Irodalomelmélet és beszédaktus-elmélet
 - 3 – 4. sz. Irányzatok a mai francia irodalomtudományban

- 1984
1. sz. Polémiák a francia forradalom előtt
 - 2 – 4. sz. Svájc népeinek irodalma – svájci irodalom?
- 1985
1. sz. A FILLM budapesti kongresszusa – A polonisztika Magyarországon
 - 2 – 4. sz. Az olasz irodalomtudomány napjainkban
- 1986
- 1 – 2. sz. A műfordítás távlatai
 - 3 – 4. sz. Színhagyományok és irodalom a mai Afrikában
- 1987
- 1 – 3. sz. A posztmodern amerikai irodalom
 4. sz. Hlebnyikov és az orosz avantgárd
- 1988
- 1 – 2. sz. Kanadai irodalmak
 - 3 – 4. sz. A stílisztika útjai és lehetőségei
- 1989
1. sz. Az empirikus irodalomtudomány elmélete
 2. sz. A felvilágosodás és nemzeti fejlődés
(A budapesti Nemzetközi Felvilágosodás Kongresszus anyagából)
 - 3 – 4. sz. A szövegkiadás új elmélete és gyakorlata: a szövegek keletkezés-kritikája
- 1990
1. sz. A mai nemzetközi folklorisztika
 - 2 – 3. sz. Irodalom és pszichoanalízis
 4. sz. A jelentésteremtő metafora
- 1991
- 1 – 2. sz. A biedermeier kora – nálunk és Európában
 - 3 – 4. sz. Hagyomány és modernizáció a mai kínai kultúrában
- 1992
1. sz. A frankofon irodalmak sajátosságai
 2. sz. Profizmus az irodalomtudományban
 - 3 – 4. sz. A Név hatalma
- 1993
1. sz. A konstruktivista irodalomtudomány
 - 2 – 3. sz. Elsikkasztott orosz irodalom
 4. sz. A mai lengyel irodalomtudomány
- 1994
- 1 – 2. sz. Az amerikai dekonstrukció
 3. sz. A kortárs olasz irodalom
 4. sz. Feminista nézőpont az irodalomtudományban
- 1995
- 1 – 2. sz. Posztszemiotika. A szubjektum-elméletek és a mai irodalomtudomány
 3. sz. A stílus diszkurzív elmélete
 4. sz. Rendszerelvű irodalomtudomány
- 1996
- 1 – 2. sz. Intertextualitás
 3. sz. Újraegyesült Németország – egységes német irodalom?
 4. sz. A posztkoloniális művelődésemélet

- 1997
- 1 – 2. sz. A félmúlt klasszikusai
 - 3. sz. Hermeneutika az orosz századelőn
 - 4. sz. A lehetséges világok poétikája
- 1998
- 1 – 2. sz. Az újhistorizmus
 - 3. sz. Kánonok a kis népek irodalmában
 - 4. sz. Textológia vagy textológiák?
- 1999
- 1 – 2. sz. A szó poétikája
 - 3. sz. Latin-amerikai irodalomelmélet
 - 4. sz. Kulturális antropológia és irodalomtudomány
- 2000
- 1 – 2. sz. A romantika tétjei
 - 3. sz. A korszakok alakzatai
 - 4. sz. (Új) filológia
- 2001
- 1. sz. Változatok a dialógusra
 - 2 – 3. sz. Dante a XX. században
 - 4. sz. Az interpretáció érvényessége
- 2002
- 1 – 2. sz. A *Poétika* újraolvasása
 - 3. sz. Autobiográfia-kutatás
 - 4. sz. A multikulturalizmus esztétikája
- 2003
- 1 – 2. sz. A minimalizmus
 - 3. sz. Mikrotörténetírás
 - 4. sz. Kísérleti irodalom
- 2004
- 1 – 2. sz. Petrarca: hermeneutika és írói személyiség
 - 3. sz. A hipertext
 - 4. sz. Wittgenstein poétikája
- 2005
- 1 – 2. sz. A kritikai kultúrakutatás
 - 3. sz. Régi az újban
 - 4. sz. Vico körei
- 2006
- 1 – 2. sz. Kritikai szubjektivizmus
 - 3. sz. Frege aktualitása
 - 4. sz. Relevancia
- 2007
- 1 – 2. sz. Alteritás, poétika, filozófia
 - 3. sz. Ökokritika
 - 4. sz. Etikai kritika
- 2008
- 1. sz. A második olvasat
 - 2 – 3. sz. A közvetítés poétikája
 - 4. sz. Az autonómia új esélyei

- 2009
- 1 – 2. sz. Eszmetörténet és irodalomtudomány
 - 3. sz. Szimbólum- és allegóriaelméletek
 - 4. sz. A jövőbelátás poétikái
- 2010
- 1 – 2. sz. Térpoétika
 - 3. sz. A félre-értelmezett futurizmus
 - 4. sz. A szerző poétikája
- 2011
- 1 – 2. sz. Testírás
 - 3. sz. Meghekkelt valóságok
 - 4. sz. Új gazdasági kritika
- 2012
- 1 – 2. sz. A Helikon repertórium 1955–2011
 - 3 – 4. sz. Boccaccio 700
- 2013
- 1. sz. Narratív design
 - 2. sz. Kognitív irodalomtudomány
 - 3. sz. Corpus alienum
 - 4. sz. Esztétika és politika
- 2014
- 1. sz. Cseh dekadencia
 - 2. sz. Funkciótörténet és kontextuális-kulturális narratológia
 - 3. sz. Az archívumok elméletei
 - 4. sz. Komparatiztikai kutatások az ezredfordulón
- 2015
- 1. sz. A kreatív írás-oktatás és a kortárs amerikai próza
 - 2. sz. Transznacionális perspektívák az irodalomtudományban
 - 3. sz. Horatius *Ars poetica*ja
 - 4. sz. Az ólomévek kultúrája és utóélete – A terrorizmus az olasz művészetekben és irodalomban
- 2016
- 1. sz. Az addikció kulturális és kritikai elméletei
 - 2. sz. Biblioterápia, irodalomterápia
 - 3. sz. Műfaj és komparatiztika
 - 4. sz. Szabadkőművesség. Új irányok a 18. századi európai maszonéria kutatásában
- 2017
- 1. sz. A százéves dada
 - 2. sz. Hálózatelmélet és irodalomtudomány
 - 3. sz. Sokarcú modernség és irodalomtörténet-írás
 - 4. sz. Fénykép és irodalom
- 2018
- 1. sz. Tzvetan Todorov, a közvetítő
 - 2. sz. Nem természetes narratológia
 - 3. sz. Próza a posztmodern (próza) után
 - 4. sz. Poszthumanizmus

2019

1. sz. Dehumanizáció: az elkövető alakja
2. sz. Térkép és irodalom
3. sz. Közép-európai komparatiztika
4. sz. Gérard Genette, a szerszámkészítő

2020

1. sz. Számítógépes irodalomtudomány

HU ISSN 0017-999X



Kiadja a Bölcsészettudományi Kutatóközpont
Irodalomtudományi Intézet
Felelős kiadó: Fodor Pál főigazgató, Kecskeméti Gábor igazgató
Nyomdai előkészítés:
BTK Történettudományi Intézet
tudományos információs témacsoport
Vezető: Kovács Éva
Tördelőszerkesztő: Zsigmondné Balázs Ildikó
A fedél és a tipográfia Benkő Anna munkája
Nyomdai munkák: Argumentum Kiadó és Nyomda Kft.
F. v.: Láng András

