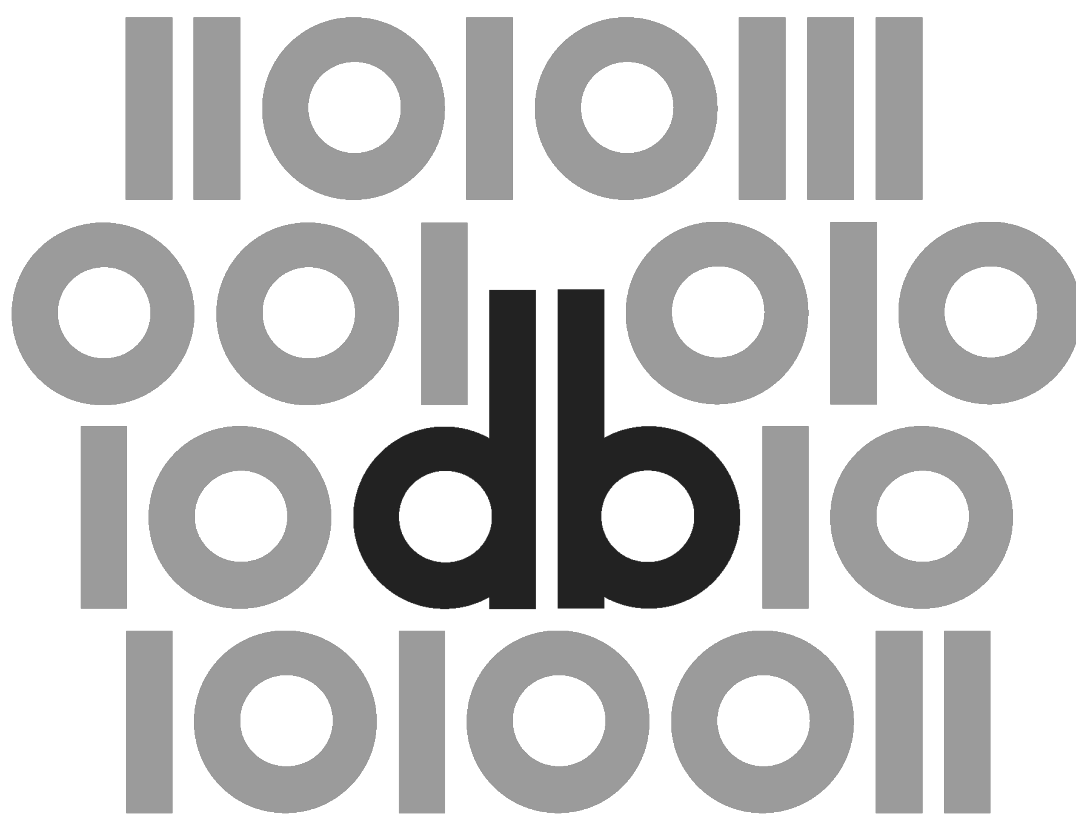


2018.01.

<DIGITÁLIS BÖLCSÉSZET>

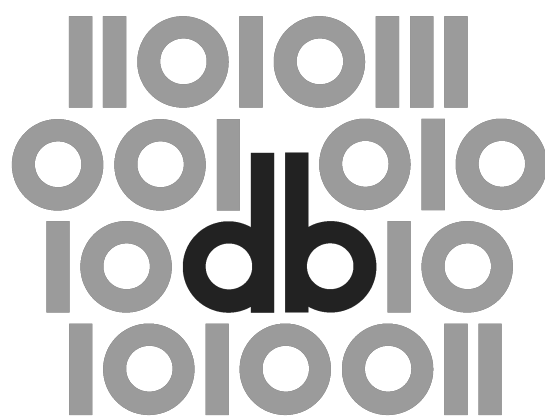


2018.01.

</DIGITÁLIS BÖLCSÉSZET>

Digitális Bölcsészet
2018., első szám

<DIGITÁLIS BÖLCSÉSZET>



2018.01.

Felelős szerkesztő:

Maróthy Szilvia

Szerkesztőbizottság:

Bartók István, Fazekas István, Golden Dániel,
Horváth Iván (a bizottság elnöke), †Orlovsky Géza,
Palkó Gábor, Pap Balázs, Sass Bálint, Seláf Levente

Szerkesztőség:

Almási Zsolt, Fodor János, Kokas Károly, †Labádi Gergely,
Parádi Andrea

Rovatvezetők:

Tanulmányok: Kiss Margit

Műhely: Péter Róbert

Kritika: Bartók Zsófia Ágnes

ISSN 2630-9696

DOI 10.31400/dh-hun.2018.1

Kiadja az ELTE BTK Régi Magyar Irodalom Tanszéke, 1088 Budapest,
Múzeum krt. 4/A.

Felelős kiadó az ELTE BTK Régi Magyar Irodalom Tanszék vezetője.

Megjelenik az Open Journal Systems (OJS) v. 3. platformon, melynek
működtetését az ELTE Egyetemi Könyvtár- és Leváltár biztosítja.

Honlap: <http://ojs.elte.hu/index.php/digitalisbolcseszett>

Email cím: dbfolyoirat@gmail.com

Tördelés: Hegedüs Béla

Grafika: Hegyi Gábor

Tartalom

Beköszöntő	7
Prószéky Gábor előszava	9
Andrew Prescott előszava	11
Szerkesztőségi köszöntő	13
Tanulmányok	15
Labádi Gergely	
Az olvasó gép: Berzsenyi Dániel versei távolról	17
Drótos László–Kokas Károly	
Webarchiválás és a történeti kutatások	35
Markó Anita	
Hálózatok a 16–17. századi album amicorumokban: Az 1500 és 1700 közötti hungarika jellegű emlékkönyvbejegyzések hálózatelemzése az <i>Inscriptiones Alborum Amicorum</i> adatbázis alapján	55
Matthew L. Jockers	
Metaadat	83
Műhely	109
Sennyey Pongrácz	
Viták és víziók a digitális bölcsészetről	111
Horváth Iván	
Digitális bölcsészet a virtuális nemzeti könyvtárban	121
Lejtovicz Katalin–Matthias Schlögl–Bernád Ágoston Zénó–Maximilian Kaiser–Peter Alexander Rumpolt	
Digitalizáció és hálózatkutatás:	
Az <i>Österreichisches Biographisches Lexikon 1815–1950</i> és az APIS-projekt	139
Cséve Anna–Fellegi Zsófia–Kómár Éva	
Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadása	
Esettanulmány	159
Biszak Sándor–Kokas Károly	
Budapest Időgép	175
Ruttkay Zsófia	
Digitális Múzeum – a MOME TechLab projektjeinek tükrében	185
Dragon Zoltán–Sebestény Csilla	
#BREW: influencer-kísérlet az Instagram újhullámos kávéközösségében	203

Kritika	217
Matthew James Driscoll and Elena Pierazzo, eds., <i>Digital Scholarly Editing: Theories and Practices</i> (2016) – Maróthy Szilvia	219
Laura Estill, Diane K. Jakacki and Michael Ullyot, eds., <i>Early Modern Studies after the Digital Turn</i> (2016) – Maczelka Csaba	223
Matthew K. Gold, ed., <i>Debates in the Digital Humanities</i> (2012); Matthew K. Gold and Lauren F. Klein, eds., <i>Debates in the Digital Humanities 2016</i> (2016) – Zámóné Kocic Larisa	233
George Bruseker, László Kovács and Franco Niccolucci, eds., „Digital Humanities.” <i>ERCIM News</i> 111 (2017) – Molnár Sándor Gyula	239
In memoriam	243
Szajbély Mihály: Búcsú Labádi Gergelytől	245

<BEKÖSZÖNTŐ>

Prószéky Gábor előszava

Van egy viszonylag új magyar kifejezés: digitális bölcsészet. A világ több részén ugyanez: Digital Humanities. A magyar változatban szereplő fogalom, a bölcsészet talán szűkebb értelmű a nemzetközileg használt humanióránál (humanities), de hazánkban az ebből a két szóból álló kifejezés vált elfogadottá. Mindez mostanában, gyakorlatilag a 21. század hajnalán történt, amikor egyre több bölcsész- és társadalomtudományi kutatási területet elért az informatika, a digitális megoldások világa. Sokaknak még persze ma sem ismerős ez a fogalom, pedig ma már a világ legnevesebb egyetemein gyakorlatilag mindenütt folyik digitálisbölcsész-képzés, és egyre-másra jelennek meg a témával foglalkozó könyvek és folyóiratok is. Érdeemes végignézni, hogy azokban a kiadványokban milyen témákról írnak, de talán érthető, ha nekünk az még fontosabb, hogy ha magyar nyelven is elindul egy ilyen kiadványsorozat, akkor abban milyen témákról lesz szó.

A hazai kutatások egyre több mindent ölelnek fel a digitális bölcsészet legkülönbözőbb ágaiból: a nyelvtechnológiától az irodalmi művek digitális kritikai kiadásain át a történelemkutatás, a muzeológia, a régészet vagy a néprajz gépesítéséig, esetleg a zenetudomány számítógépes kutatásáig, de még folytathatnánk számos, újnak tűnő szakterület nevének felsorolásával. A közelmúlt egyik érdekes témájú doktori védésén az opponens épp azt fogalmazta meg, hogy „az utóbbi években egyre több olyan kutatás készül, s ezek nyomán egyre több olyan doktori értekezés lát napvilágot a bölcsészet és társtudományai területén, mely afféle sajátos, «digitális ugarfeltörésként» is értelmezhető. Ezek a művek olyan területekre terjesztik ugyanis ki a digitális kultúra érvényességét, melyeket addig e kultúrával ellentétesnek gondolt a közvélemény, vagy éppen az adott terület «őslakossága». A mára gyakorlatilag bevett «digitális bölcsészet» kifejezés például néhány éve Magyarországon még afféle «fából vaskarikának» számított, hiszen a digitális kultúrára jellemzőnek vélt kvantifikáló, tárgyiasító, áruképző szemléletnél kevés dolog áll távolabb attól, amit a bölcsészetről gondoltak a bölcsészek.”¹

Ami nagyon lelkesítő ezzel az új kutatási területtel kapcsolatban, az éppen az, hogy ma már mindenfajta kutatásba jól integrálható néhány olyan megoldás, melyet az informatika eszközeivel lehet csak az adott területen létrehozni. Akik ilyen kutatásokat végeznek, azok számára a bölcsész szakterület és az informatikai technikák integrálása szinte automatikusan történik. Nyilván a „digitális bölcsészet” kifejezés – ahogy feljebb is utaltunk rá – ma még sok helyen újdonságnak számít, vagyis az egyes bölcsészterületek informatikával ötvözött különféle kutatásait még joggal soroljuk a digitális bölcsészet ernyőfogalom alá. Eljön azonban az az idő, és talán nem is a túl távoli jövőben, amikor a „digitális” jelző lassan eltűnik, és amint más területeken is, a bölcsészet világában is a digitális megoldások természetes kutatási eszköznek számítanak majd. Ekkor a bölcsészet fogalma már automatikusan magában foglalja a digitális bölcsészetet, természetessé válik az informatika használata, amihez már nem fog kelleni önálló név.

¹ Nagy Károly Zsolt opponensi véleménye Bolya Máttyás „Információelmélet és népzene-kutatás” című doktori értekezéséről (2018).

De most még kell önálló név, ezért ennek a diszciplínának a bemutatására, az ezen a területen folyó kutatások ismertetésére indul most útnak a *Digitális Bölcsészet* című elektronikus folyóirat.

Prószéky Gábor

Andrew Prescott előszava

It is a great honour to be asked to provide a foreword to the first Hungarian journal on Digital Humanities, *Digitális Bölcsészet*. The use of computational techniques by Hungarian scholars has a long and distinguished history. Indeed, it was at the University of Debrecen in July 1998 that the Association of Literary and Linguistic Computing inaugurated the most prestigious award for lifetime achievement in digital humanities, the Busa Prize, named after Father Roberto Busa, whose use of computers to create the *Index Thomisticus* is regarded as marking the foundation of digital humanities.

Father Busa was himself the first recipient of the Busa Prize and Busa's lecture at the University of Debrecen on the occasion of this award was characteristically playful and visionary. Busa declared that:

Computers, which God wanted His children to give themselves, mark the beginning of a new Era [...] . They should mark also, and promise, a new era in language knowledge, i.e. an enhancing of scientific quality, a leap of dimensions, a deepening, an enlarging of information about human expression.¹

For Busa, language and knowledge were profoundly intertwined, and computers, by transforming our ability to analyse language, also offered the prospect of an epistemological revolution. In an intriguing sally, Busa urged that computational methods should not be restricted to literary texts, but extended to film and sound:

Would it not be reasonable to process, also and in addition, e.g. the academic manuals or courses of any discipline from sciences to history, from journalism to film-making? Consider it.²

Busa described the commercial computing of 1998 as an ocean. If he were to return to Debrecen now, he would doubtless feel that the use of computers in humanities has also become an ocean. This inaugural issue of *Digitális Bölcsészet* illustrates the range and variety of current digital humanities and shows how Hungarian scholars active in this field match their peers elsewhere in the scale, range and inventiveness of their work.

Running through much of the humanities is a preoccupation with the analysis of primary cultural materials, ranging from novels, poems, newspapers, historical records and material artefacts to paintings, film, sculpture, music, television and increasingly born-digital media. What unites the digital humanities is a belief that, following Busa's lead, computational methods of different types can shed new light on the nature of these primary materials. These computational methods as they draw together these primary materials in new configurations can create new cross-disciplinary dialogues and, as Busa hoped, foster new forms of epistemological understanding.

¹ R. Busa, „Picture a Man: Busa Award Lecture, Debrecen, Hungary, 6 July 1998,” *Literary and Linguistic Computing* 14, no. 1. (1999): 5–9.

² Busa, „Picture a Man,” 8.

The essays in this volume reflect the extraordinary range of methods deployed by practitioners of the digital humanities. Historical and literary materials are reimaged as networks or geospatial entities. Distant readings subject some groups of texts to high-level analysis, while other projects exploit the ability of computers to grind text very finely. Formal computational techniques requiring collaboration with computer sciences jostle with social media. Scholars grapple with the new research techniques required to make use of born-digital materials such as web archives.

This can all be seen as very much in the spirit of Busa. In his 1998 Debrecen lecture, Busa urged humanities scholars to make use of all the possibilities of information technology. He cited an epigram of Plautus, translated as "Tis as you'd turn a stream upon your field; which if you do not, it will all run waste into the sea."³ In other words, if we don't use some of the water from the digital floods which surround us to cultivate our field, it will simply run into the ocean and be wasted.

In contemplating the current state of digital humanities, the variety of techniques and subject matter may seem bewildering. But the lack of rigid boundaries and set methods is a good thing. We are all now aware of how vast and varied the oceans of the digital world are. Digital humanities represent part of those oceans. The waters may be choppy, the currents bewildering and the navigational charts confusing, but the voyages open new horizons and point the way to new worlds.

This interconnection with the wider digital world is important. As our scholarship becomes increasingly absorbed in a digital environment, it is vital that we retain critical perspectives. The way in which we study humanities in a digital environment is profoundly different from humanities conducted in an old-fashioned library or archive. This is why digital humanities has in recent years, particularly in the United States, strongly intersected with wider critical debates about the cultural and social biases of knowledge.

These cultural discussions are important and vital because they help construct a critical understanding of the digital world where so much of our life nowadays take place. As recent scandals over abuse of data illustrate, the need to develop strong critical and ethical frameworks for our digital engagement is becoming increasingly pressing. When so much of our society and culture depend on the internet, the health of the internet is vital to us all.⁴

Digital humanities has an important role to play in creating a healthy digital culture and society. This first issue of *Digitális Bölcsészet* demonstrates that the digital humanities in Hungary has an important and distinctive contribution to make to these endeavours. I wish that this journal, which following Busa's metaphor I might compare to a small stream, will grow into a river, feeding into a mighty river of European digital humanities and thence flowing into a wider ocean of critical digital knowledge. I wish those undertaking this voyage every success.

Andrew Prescott

³ Busa, „Picture a Man,” 5.

⁴ <https://internethealthreport.org/>

Szerkesztőségi köszöntő

Örömmel bocsátjuk útjára a magyar *Digitális Bölcsészet* tudományos folyóirat első számát. Mit is jelent azonban az, hogy az „első” számot, és mit értünk azon, hogy „örömmel”?

Az „első” mint sorszám az újdonságot és a sorozatszerűséget jelöli. Bár a digitális kutatások, a számítógép használata Magyarországon a bölcsészettudományok és társadalomtudományok területén az 1970-es évekre tekint vissza, a folyóirat most mégis fontosnak tartja rögzíteni azt a tényt, hogy a digitális bölcsészet mint diszciplína, illetve mint elméletek és gyakorlatok összessége Magyarországon és magyarul is létezik. Az „első” továbbá arra is felhívja a figyelmet, hogy a folyóirat nemcsak hogy létezik, de további számok megjelentetését is tervezzük, sőt mire ez a szám napvilágot lát, a szerkesztőbizottság már a következő számon fog dolgozni. Az „első” tehát egy szekvencia kezdetét jelöli, egy sorozatot, amely elindult, és folytatódni akar.

Az útra bocsátás örömeiben benne foglaltatik az, hogy érdemes folytatni a munkát. Érdemes, hiszen az első szám cikkeit tekintve a digitális bölcsészeti tevékenységeknek sokszínűségét egy hagyományosnak tekinthető médiumon hagyományosnak mondható szerkezetben tárja az olvasó elé. A digitális bölcsészet egyszerre érdeklődik az elméleti reflexió és a gyakorlati megvalósulás iránt, éppen ezért találhatunk hosszabb és rövidebb tanulmányokat is. Ennek megfelelően a folyóirat és a jelen szám is három nagyobb csoportba osztja a cikkeket: „Tanulmányok” (rovatszerkesztő: Kiss Margit), amelyek hosszabb lélegzetű, elméleti megfontolásokat tárnak az olvasók elé, „Műhely” (rovatszerkesztő: Péter Róbert), ahol projektbeszámolókat és alkalmazáselemzéseket olvashatunk. A „Kritika” rovat zárja a sort (rovatszerkesztő: Bartók Zsófia Ágnes), ahol rövidebb könyv-, honlap- és folyóiratszám ismertetéseket, valamint recenziókat találhatunk.

A „Tanulmányok” rovatban négy cikk foglalkozik a digitális bölcsészet elméleti jellegű kérdéseivel: Labádi a gépi szövegelemzéssel foglalkozik, Drótos és Kokas a webarchiválás és a történettudomány kapcsolatát elemzi, míg Markó a digitális hálózatkutatás eszköztárával analizálja az 1500–1700 között keletkezett *inscriptiók* alapján a korabeli tudós hálózatokat, végül Jockers (Labádi fordításában) bibliográfiai és demográfiai metaadatok elemzésével törekszik egy adott irodalmi narratívát újraértékelni. A „Műhely” rovatban Ruttkay a digitális múzeumi projektekről, Biszak és Kokas a Budapest-időgépről, Bernád, Kaiser, Lejtovitz, Rumpolt és Schlögl az APIS-projektről, Cséve, Fellegi és Kómár Móricz levelezésének digitális kritikai kiadásáról, Sennyey a digitális bölcsészetről vitákról, Horváth egy virtuális könyvtárra vonatkozó hipotetikus álláspontokról, míg Dragon és Sebestény egy Instagramon folytatott kísérletről számol be. A „Kritika” rovatban Maróthy, Koczicz, Maczelka és Molnár digitális bölcsészethez kapcsolódó monográfiákat, valamint egy folyóirat különszámot szemlél.

Az első szám szerkezetét tekintve bár hagyományosnak is mondható, talán annyiban mégis formabontó, hogy nemcsak egy diszciplína megszületése feletti örömről árulkodik, hanem az esendőségről, a hiányról is, ahogy az egyébként a bináris kódnak a jellegzetessége: 1 0, az igen-nem, lét-nemlét. Az esendőségről, amennyiben nemcsak a születés öröme, hanem a halál fájdalma is benne foglaltatik, hiszen kedves kollégánk,

barátunk, Labádi Gergely, a magyar digitális bölcsészet egyik kiemelkedő alakja 2017 szeptemberében távozott el közülünk, és az első számban egy rá, szerkesztőtársunkra emlékező írást is közlünk. A kezdetekkor rövid ideig a szerkesztőbizottság tagja volt az ugyancsak 2017-ben váratlanul elhunyt Orlovsky Géza is, aki a folyóirat indulásakor (az egykor a Bölcsészeti Informatika Önálló Programnak is helyt adó) ELTE Régi Magyar Irodalom Tanszékének vezetőjeként állt mellettünk.

Továbbá annyiban is formabontónak értékelhetjük a jelen folyóiratszámot, hogy a születés és halál közötti területre is utal, ti. a történelemre. A történelemre, a jelenvaló történelemre tereljük a figyelmet, amikor a magyar digitális irodalomtudomány egy másik ikonikus alakjának, Horváth Ivánnak közöljük egy régebbi, kibővített és szerkesztett konferencia-előadását. Ezzel egyúttal jelezni szeretnénk, hogy a tervezett elavulás, amely a digitális világban meglehetősen rövid időt foglal magában, a digitális reflexióra esetenként mégsem érvényes.

Az öröm arra is vonatkozik, hogy egy folyóirat nemcsak azért érdekes, mert sokszínű és emlékező, hanem azért is, mert már maga egyfajta infrastruktúrát jelent. Egy folyóirat nemcsak azért lehet jelentős, különösen érvényes ez a *Digitális Bölcsészet* címűre, mert az érdeklődő, a tudós megismerkedhet mások gondolataival, rácsodálkozhat mások projektjeire, hanem azért is, mert a folyóirat mindezek által fórumot teremt a diszkusszióra. Arra a diszkusszióra, amely a tudomány lételeme, ahol problémákat lehet felvetni és elemezni, ahol ezeken tovább lehet gondolkodni, ahol vissza lehet kérdezni, ahol kutatók közölnek és hallgatnak, tudatnak és beszédbe elegyednek egymással. Más szóval a folyóirat, mondhatnánk „hálózatot”, de mondjuk inkább „tudós közösséget”, netán „tudós barátságokat” hoz létre. A folyóirat azáltal, hogy befogad cikkeket, hogy bírálóknak adja őket, hogy azok reflektálnak, majd a reflexióikra válaszolnak a szerzők, ahogy ebbe a szerkesztők is bele-beleszólnak, mind azt segítik elő, hogy ne csak egyszólamú, független monológok, hanem tudós beszélgetések, tudós közösségek alakulhassanak ki. Infrastruktúra annyiban is a folyóirat, amennyiben a nyelv is egyfajta környezet, amely lehetővé teszi a diskurzust. A folyóirat tehát egyúttal a digitális bölcsészet magyar nyelvű tudományos diskurzusához, terminológiájának kialakításához igyekszik hozzájárulni.

Örömmel bocsátjuk hát útjára az első számot, hogy további közösséget, tudományos közösséget alkothasson immár a szerkesztőktől függetlenül, önállóan is.

A szerkesztőség nevében:

Almási Zsolt

Köszönetnyilvánítás

A *Digitális Bölcsészet* folyóirat létrehozása az alábbi intézmények, projektek támogatásával valósult meg: ELTE Régi Magyar Irodalom Tanszék, ELTE Egyetemi Könyvtár és Levéltár, ELTE Dékáni Titkárság, EFOP 3.6.1–16–2016–00008 azonosítójú, EU társfinanszírozású projekt (SZTE), SZTE Klebelsberg Könyvtár.

Köszönjük továbbá Garamvölgyi László, Hegedüs Béla, Hegyi Gábor, Matthew L. Jockers és Virág Gabriella közreműködését!

A szerkesztők

<TANULMÁNYOK>

Labádi Gergely †

Szegedi Tudományegyetem, Magyar Irodalmi Tanszék

Az olvasó gép: Berzsenyi Dániel versei távolról*

Tanulmányom helyzetfelmérés és kísérlet. Az érdekel, a rendelkezésre álló magyar nyelvi számítógépes elemzőkkel és statisztikai programokkal lehetséges-e a Berzsenyi-versek csoportosításával kapcsolatos, a szoros olvasás révén felmerült hipotézist igazolni, illetve meg lehet-e válaszolni olyan, Berzsenyi hangzóelméleti és -használati gyakorlatát illető kérdéseket, amelyeket szoros olvasással nyilvánvalóan nem lehetne, emellett bemutatok egy lehetséges elemzési módszert, amely egyes fogalmak felbukkanását, eloszlását és összefüggéseit kutatja. Olyan belső kapcsolatokat, struktúrákat tárhatunk fel, illetve olyan modelleket, törvényszerűségeket állapíthatunk meg, amelyek a szoros olvasás révén egyszerűen nem állnának rendelkezésünkre. Egy olyan szemléletmódról és módszertanról van tehát szó, amely korábban elképzelhetetlen mértékű kontextualizációt tesz lehetővé, amely az egyedi szövegek környezetének, az egyes szövegeket meghatározó irodalmi, társadalmi, kulturális trendeknek az azonosítását, elemzését ígéri: ez a történeti és stilisztikai kérdéseknél egyértelmű előnnyel rendelkezik a szoros olvasással szemben.

Kulcsszavak:

Berzsenyi Dániel, R, Magyarlanc, tartalomelemzés, távoli olvasás



1. Géppel mérni

Tanulmányom helyzetfelmérés és kísérlet. Az elmúlt néhány évben a klasszikus irodalomtörténeti, irodalomtudományos kérdések mellett – mint oly sokunkat – a digitális bölcsészet elméleti és gyakorlati kérdései, problémái is foglalkoztattak.¹ A számítógép használata immár végérvényesen és egyértelműen nem pusztán eszközprobléma, mint

* A kutatást az EFOP-3.6.1-16-2016-00008 azonosítójú, EU társfinanszírozású projekt támogatta (Intelligens élettudományi technológiák, módszertanok, alkalmazások fejlesztése és innovatív folyamatok, szolgáltatások kialakítása a szegedi tudásbázisra építve).

¹ Labádi Gergely, „A filológiai tudás formái,” in *Textológia – filológia – értelmezés: Klasszikus magyar irodalom*, szerk. Czifra Mariann és Szilágyi Márton (Debrecen: Debreceni Egyetemi Kiadó, 2014), 173–190; Labádi Gergely, „A magyar regény adatbázisa,” *Acta Historiae Litterarum Hungaricarum* 32 (2016): 11–30. Szintén ennek az érdeklődésnek jegyében szerveztük meg Kokas Károllyal és Péter Róberttel közösen 2015 októberében a *Digitális bölcsészet Szegeden* című workshopot (hozzáférés: 2018.05.23, <http://digibolcsesz.ek.szte.hu/>).

aminek az 1970-es években esetleg látszott.² A '90-es évek közepének felkiáltása („The Information Society is now upon us!”) ugyan jelezheti egy buzzword aktuális tündöklését, de mint tudjuk, ennél éppenséggel többről van szó. Az információs társadalom a modernitás uralkodó médiumának, a könyv pozícióinak megroppantásával kihívást jelent a humántudományok számára mind a gyakorlat, mind az elmélet vonatkozásában. Hivatkozhatunk akár a mindennapokra is, arra, hogy – legalábbis a nyugati tudományosságban – a *Digital Humanities* a 2000-es évek második felére szervezetileg, intézményileg végleg és egyértelműen áttört,³ azaz vannak képzések, folyóiratok, konferenciák vagy minimum konferenciaszekciók, a bölcsészállásoknál rendszerint elvárás valamiféle *DH*-képeség vagy -gyakorlat. A számítógépek használata azonban sokáig csak a filológia klasszikus gyakorlatainak könnyítését célozta, úgy tettünk, mintha valójában semmi sem történt volna.⁴ Mellőzve most a humántudományok gyakorlatának korábban már tárgyalt kérdéseit,⁵ pusztán az a módszertani megközelítés, amelyet ki *macroanalysis*ként, ki *distant*, ki *machine reading*ként, ki pedig *algorithmic criticism*ként nevez meg, sürget bennünket, hogy újragondoljuk eddigi, *close reading*ként emlegetett gyakorlatunkat és annak eredményeit.

Az irodalomtudományban eddig a „bizonyítékok” gyűjtése alapvetően szubjektív megfigyeléseken alapult, a belőlük levont általánosítás érvénye pedig a minta reprezentativitásán múlt. Matthew Jockers példája Ian Wattnak a regény felemelkedéséről szóló klasszikus munkája, amely mindössze Daniel Defoe, Samuel Richardson és Henry Fielding munkái alapján készült, jöllehet a szóba jöhető szerzők és szövegek köre nagyságrendekkel nagyobb. Adódik a kérdés, ha ezt a több ezer további szöveget el tudnánk „olvasni,” más történetet mondanánk? Eddig ezt a kérdést esélyünk sem volt megválaszolni, de most már megvannak az eszközeink.

A számítógép-használat adatok gyűjtésére valójában mégsem eszközkérdés. Hiszen – durva leegyszerűsítéssel – a statisztikák irodalmi használatának elméleti és módszertani alapjai több évtizedesek, a szerzőattribúció elméletei éppenséggel még régebbiek. Tehát nem pusztán eszközkérdés: mostanra „egy olyan fordulóponthoz, eseményhorizonthoz értünk, ahol már elegendő szöveget és szakirodalmat kódoltunk ahhoz, hogy lehetővé tegye számunkra, sőt hogy kényszerítsen bennünket, új kérdéseket tegyünk fel az irodalomról és az irodalmi adatról.”⁶ Matthew Jockers elbeszélésében 2008 a fordulópont, és bár nem reflektál rá, de megállapítása természetesen csak az angol nyelvű szöveghagyományra vonatkozik. A német kultúra kapcsán Fotis

² Lásd a Jockers-kritikámban (Labádi Gergely, „Matthew Jockers, *Macroanalysis: Digital Methods and Literary History*,” *Irodalomtörténet* 97, 4. sz. (2016): 496–500) idézett tanulmányokat: Martin L. West, *Szövegkritika és szövegkiadás*, ford. Bolonyai Gábor (Budapest: Typotex Kiadó, 1999 [1973]); Voigt Vilmos, „Számítógépes ritmuselemzési kísérlet,” *Irodalomtörténeti Közlemények* 76, 2. sz. (1972): 203–211.

³ Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013), 11–23.

⁴ „Der Einsatz von EDV in der editorischen Arbeit grundsätzlich eine Arbeitserleichterung und eine weitaus grössere Präzision in den zu erzielenden Ergebnissen mit sich bringt.” Bodo Plachta, *Editionswissenschaft: Eine Einführung in Methode und Praxis der Edition neuerer Texte* (Stuttgart: Reclam, 1997), 131.

⁵ Labádi, „A filológiai tudás,” 173–190.

⁶ Jockers, *Macroanalysis*, 4.

Jannidis és Gerhard Lauer szerint 2011, a TextGrid repozitóriumának megindulása óta áll fenn a kvantitatív vizsgálatok alapfeltétele, a megfelelő mennyiségű és minőségű digitalizált szöveg.⁷ Mi a helyzet a magyar nyelven írott hagyománnyal? Az Arcanum szolgáltatásában az elmúlt 200 év magyar tudományának több mint 11 millió oldalnyi anyaga van digitalizálva, a Magyar Elektronikus Könyvtár (MEK) 16000 dokumentum fölött jár, a Digitális Irodalmi Akadémiában (DIA) pedig több mint 80 kortárs életmű férhető hozzá. Az Országos Széchényi Könyvtár ELDORADO szolgáltatása a még nem digitalizált források kérdését próbálja, kívánja rendezni. Érdemes tehát nekünk is elgondolkodni, új kérdéseket feltenni, még ha jelenlegi formájában az előbb felsoroltak közül nem is minden alkalmas kvantitatív vizsgálatok lefolytatására – a kétrétegű PDF például korántsem a digitális mint olyan kvintesszenciája; a DIA anyaga nem könnyen hozzáférhető, a MEK szövegeinek megbízhatósága pedig problematikus, azaz a magyarországi digitalizált szövegek nem egyenletes minőségűek, és nem egységes elvek szerint készültek.

Az új megközelítés kiindulópontja, bárhogya is nevezi magát, hogy az egyedi szövegtől eltávolodva olyan redukciós és absztrakciós eszközöket nyerünk, amelyek a tudás speciális, a konkrét szöveg realitása felől nem észlelhető formáját kínálják. Olyan belső kapcsolatokat, struktúrákat tárhatunk fel, illetve olyan modelleket, törvényszerűségeket állapíthatunk meg, amelyek a szoros olvasás révén egyszerűen nem állnának rendelkezésünkre. Egy olyan szemléletmódról és módszertanról van tehát szó, amely korábban elképzelhetetlen mértékű kontextualizációt tesz lehetővé, amely az egyedi szövegek környezetének, az egyes szövegeket meghatározó irodalmi, társadalmi, kulturális trendeknek az azonosítását, elemzését ígéri: ez a történeti és stilisztikai kérdéseknél egyértelmű előnnyel rendelkezik a szoros olvasással szemben. Ugyanakkor nagy mennyiségű adat számítógépes elemzése egyszerűen csak egy alternatív módszer adatokat gyűjteni, hipotéziseket felállítani, ám ezeket ugyanúgy kritikával kell kezelni, és értelmezni szükséges.

A „nem-olvasás” – ahogy Franco Moretti saját gyakorlatát provokatívan nevezi⁸ – tehát használható korábbi hipotézisek ellenőrzésére, új hipotézisek felállítására, korábban fel nem ismert, lehetséges kutatási területek körülhatárolására. Mivel Moretti és Jockers könyvéről írott recenzióimban már utaltam néhány konkrét példára, most inkább olyan területeket neveznék meg, amelyek megoldására eddig nemigen volt lehetőségünk, ugyanakkor a számítógépek használata révén lehetséges, vagy legalábbis erős érveket kaphat egyik vagy másik hipotézis, tehát lépéseket tehetünk a megoldás érdekében. Ha tétellel bíró vizsgálatot keresünk, akkor rögtön adódnak az eldöntetlen szerzőségű szövegek. Ilyen akad bőséggel a magyar irodalomban, és a kanonikus szövegek között például a *Fanni hagyományai* szerzősége az egyik leg-

⁷ Fotis Jannidis and Gerhard Lauer, „Burrows’s Delta and Its Use in German Literary History” in *Distant Readings: Topologies of German Culture in the Long Nineteenth Century*, eds. Matt Erlin and Lynne Tatlock (Rochester: Camden House, 2014), 30.

⁸ „What does it mean, studying world literature? How do we do it? I work on West European narrative between 1790 and 1930, and already feel like a charlatan outside of Britain or France. World literature? Many people have read more and better than I have, of course, but still, we are talking of hundreds of languages and literatures here. Reading more seems hardly to be the solution. [...] Reading more is always a good thing, but not the solution.” Franco Moretti, „Conjectures on World Literature,” *New Left Review* 1 (2000. jan.-febr.): 55.

nagyobb filológiai rejtélyek egyike. Az értelmezéstörténet sokkal inkább az aktuális irodalompolitikai kívánalmak szerint alakult, semmint valamiféle ellenőrizhető filológiai, nyelvi módszertan alapján. Pedig Szinnyei Ferenc egyszer már megpróbálta nyelvészeti módszerekkel megoldani a kérdést, de a szakirodalom nem fogadja el érveit – még ha egybeesik is a mostani konszenzussal.⁹ Ennek azonban nemcsak a kiindulópont bizonytalansága az oka, maga a követett módszertan sem meggyőző. A szerzőattribúció során ugyanis tipikusan nem egy-egy tartalommal bíró kifejezést kell vizsgálni, hanem a leggyakoribb szavakat (*most frequent words*), amelyek között a funkciószavak (például névmások, névelők, viszonyszók) dominálnak – Jockers szerint még az írásjelhasználat is árulkodó –, illetve e szavaknak a szöveg egészéhez viszonyított arányát kell nézni.¹⁰ Szinnyei viszont már egy-egy tartalommal bíró szó felbukkanását döntőnek tartja. Időnként persze maga is elismeri, hogy némelyik kifejezés elég népszerű a kor érzékeny prózájában, tehát szerzőattribúcióra kevésbé alkalmas – mint például a *csendes*, az *édes* vagy a *kedves* –, vagy hogy valójában a szöveg „numerozitása,” a felkiáltások szintén az érzékenység stílusjellemzői. És akkor még nem beszéltünk a szerkesztői négykezesek lehetőségéről sem – ami a korszakból az *Egyéni és eszményi* attribúciójának is kérdése egyébként –, amelyet viszont a számítógépes módszerek már kezelni tudnak.¹¹

Szintén szerzőattribúciós kérdés, ugyanakkor jóval hevesebb indulatokat váltott ki még szakmai körökben is Kosztolányi Dezső név nélküli cikkeinek kérdése, különösen az *Új Nemzedék*ben vállalt szerepének története.¹² A vitához új adatokkal járulnának hozzá a körületekintően használt, előtte alaposan, semleges szövegeken kipróbálva „belőtt” szerzőattribúciós vizsgálatok. A századelő újságírói gyakorlatának filológiai kérdéseit az irodalomtörténeti szakma szintén nyelvi, stiláris eszközökkel kívánta megoldani, ugyanakkor a követett módszertan, ha talán közelebb is áll a szerzőattribúciós kutatások mai állásához, mint Szinnyei javaslata, nem gépi eszközöket használ,¹³ ezért az eredményei kétségesek, szubjektívabbak. Sok bennük a megérzés: Péter László nem véletlenül említi fent hivatkozott tanulmánya utolsó bekezdésében, hogy számítógéppel kellene folytatni a vizsgálatokat, többet és többfélét lehetne elemezni.

Az irodalmi szövegekkel kapcsolatban vannak ugyanakkor más jellegű, az írói nyelvhasználathoz (is) kötődő, de sosem igazolt kijelentések. Mikszáth elbeszélőmű-

⁹ Szinnyei Ferenc, *Kármán József és az Uránia névtelenjei* (Budapest: MTA, 1924). Vö. Szilágyi Márton, *Kármán József és Pajor Gáspár Urániája* (Debrecen: Kossuth Egyetemi Kiadó, 1998), 80.

¹⁰ Jockers is foglalkozik a kérdéssel, de nagyhatású volt Burrows tanulmánya: John Burrows, „Questions of Authorship: Attribution and Beyond,” *Computers and the Humanities* 37 (2003): 5–32, <https://doi.org/10.1023/A:1021814530952>. Informatív történeti összefoglalásként pedig: Harold Love, *Attributing Authorship: An Introduction* (Cambridge: Cambridge University Press, 2002), <https://doi.org/10.1017/CB09780511483165.001>.

¹¹ Brian Vickers, *Shakespeare, Co-Author: A Historical Study of Five Collaborative Plays* (Oxford: Oxford University Press, 2002).

¹² Számos cikk foglalkozik a kérdéssel, csak egy összefoglalót idézek, amely utal a vita több résztvevőjének írására: Lengyel András, „Egy s más az Új Nemzedék Pardon rovatáról,” *Kalligram* 19, 6. sz. (2010): 89–99. Köszönöm Bíró-Balogh Tamásnak a kérdés történetéről adott felvilágosítását.

¹³ Péter László, „Juhász Gyula névtelen cikkeinek felismerése stílussajátságai alapján” in *Jelentés és stilisztika: A Magyar Nyelvészek 2. Nemzetközi Kongresszusának előadásai*, szerk. Imre Samu, Szathmári István és Szüts László (Budapest: Akadémiai Kiadó, 1974), 454–457; Lengyel András, „Kosztolányi-dubiózák,” *Forrás* 38, 11. sz. (2006): 91–113.

vészete, különösen a *Tót atyafiak* és *A jó palócok* kapcsán már az első megjelenés óta gyakran elhangzik az előbeszédszerűség kategóriája. Az értelmezésekben azonban ennek kifejtése általában meglehetősen homályos marad. A Tahin Szabolcs által összegyűjtött kijelentések jól jelzik, mennyire nehéz megfogni e sajátosságát a novelláknak: „Mintha csakugyan eleven beszéd csengene fülembe olvasásakor...” „Mintha nem is volna köztünk az a nagy távolság, mely író és olvasót elválasztja...” „Mintha minden szó egyenesen az ajkáról lebbent volna a papírosra...”¹⁴ Ugyanakkor bármilyen homályosak is az idézetek, az egyértelmű, hogy az előbeszéd az írásbeliséggel van szembeállítva. Ebben az esetben pedig lennie kell mérhető különbségeknek – vannak is.¹⁵ Az előbeszédszerűség narratológiai elemzései, amelyek a novellák elbeszélőjéhez, elbeszélői tudatához kapcsolva hozzák fel, természetesen meggyőzők, és ettől a vizsgálattól függetlenül, bár a továbblépés útját mutatják a különböző „hangok” stiláris sajátosságainak vizsgálatára figyelmeztetve. Herczeg Ferenc egyes regényeinek szókincsével kapcsolatban is elhangzott a szegényes minősítés – ez is mérhető adat.¹⁶

Heves reakciókat vált ki, amikor egy-egy klasszikust modernizálnak. Az átdolgozók amellet érvelnek, hogy így közelebb hozzák a szöveget a mai diákokhoz, hogy majd így kedvet kapnak az eredetihez, az ellenzők viszont – már-már a morális pánik jeleit mutatva – azt állítják, hogy csak elcsökevényesítjük a diákokat, ha nem az eredetivel szembesülnek. A vitát megoldani persze nem lehet, de a számítógépes szövegvizsgálatok abban segíthetnek, hogy megválaszoljuk, valójában mely tulajdonságai tűntek el a régi szövegnek, megőrzött-e valamit az eredeti nyelvi struktúrákból, mit domborít ki az új verzió, mennyiben veszi át az átdolgozó szövegeinek nyelvi sajátosságait.¹⁷ A mérhető sajátságok tárháza széles – a mondatok összetettségétől kezdve a szófaji arányokon át az átlagos szóhosszúságig –, és valójában hozzásegíthet bennünket az eredeti néhány, korábban talán nem is észlelt sajátosságának fölismeréséhez. Valamint a sikeres/nem sikeres írók közti különbségek is mérhetőek – még ha a mérések nem is magyaráznak meg mindent.¹⁸

2. Verseket mérni

Az eddig hivatkozott tanulmányok, a Shakespeare-szakirodalom kivételével, mind prózai szövegekkel foglalkoznak, de természetesen vannak olyan tanulmányok is, amelyek verseket vizsgálnak. Ezek azonban Berzsenyi esetében kevésbé segítenek. Az angol nyelvű költészet sikeres, antológiákban gyakran publikált darabjait vizsgáló tanulmány eredményei – az első sor jellemzően kevesebb szótagból áll, mint „ismeretlen” kortársának verse, és jellemzően monoszillabák alkotják, továbbá a sikeres költeményeknek egyszerűbb a szótára, rövidebbek a szavai – nem túl érdekesek, illetve

¹⁴ Tahin Szabolcs, „»Előbeszédszerűség« Mikszáth prózájában,” *Tisztaíj* 57, 11. sz. (2003): 53–71.

¹⁵ Labádi Gergely, *Géppel mért irodalom: a mikszáthi előbeszédszerűség*, kézirat, 2017.

¹⁶ Tóth Mihály, *Herczeg Ferenc írói szókincsének vizsgálata*, kézirat, 2017.

¹⁷ A több klasszikust, például az *Egri csillagokat*, *A kőszívű ember fiait* átdolgozó Nógrádi Gergely maga is író.

¹⁸ Jodie Archer and Matthew L. Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel* (New York: St. Martins Press, 2016).

egy-egy költői korpusz vizsgálatára kevésbé alkalmasak.¹⁹ Van ugyan izgalmasabb eredményeket felmutató számítógépes kutatás is, amely az angol nyelvű amatőr és professzionális költők versei közötti különbségeket keresve arra jutott, hogy elsősorban nem a versírás megtanulható és könnyen mérhető sajátosságai jelzik a különbséget (mint a rímek vagy a prozódia), hanem a használt kifejezések konkrétsága.²⁰ Ez azonban a megfelelő háttéranyag nélkül, ti. a magyar szavak konkrétsági fokát jelző szótár nélkül nálunk nem alkalmazható – a szerzők és Simon Eszter jóvoltából hozzám eljutott kéziratos cikk sajnos nem túl nagy (296 főnév) korpuszt dolgoz fel²¹ –, jóllehet a megállapítás egybeesik a 19. század második felének Hász-Fehér Katalin által feltárt kritikai diskurzusával.²² A lengyel költészet grammatikai rímeinek, a grammatikairím-használat történetének statisztikai vizsgálata ugyan a magyar anyag kapcsán is ígéretes, ám megfelelő háttéranyag nélkül egy-egy költői korpusz elemzésére nem lehet használni.²³

A kísérletet ugyanakkor támogatja, hogy versek statisztikai jellegű vizsgálatának magyar szakirodalma, kutatástörténete is van, amely pontosan megnevezi, milyen adatokat milyen intencióval mér. Zsilka Tibor egyik tanulmánya például kifejezetten a magyar költői nyelv statisztikai vizsgálatának, a vizsgálat szabályainak kialakítását tűzte ki célul azzal a szándékkal, hogy „egy oly korban, amikor garmadával jelennek meg a különböző színvonalú s tartalmú verskötegek és egyéb írásos művek,” legyenek olyan „egzakt módszerek,” amelyekkel – túl „az intuíció”-n – megítélhetők.²⁴ Bár számos izgalmas megállapítást tartalmaznak Zsilka vizsgálatai, azzal nem tudok egyetérteni, hogy egy szerző azért volna jobb költő, mint a másik, mert a jeltípusok (szótári szavak, a képletben: V) és a tényleges szóalakok (N) aránya (Type Token Ratio) nála a magasabb, azaz választékosabb, „gazdagabb” a szókincse, mint költőtársáé, vagy esetleg kisebbek egyes szófajok ismétlődési arányai, vagy mert több ige használ, mint a másik, tehát „stílusa dinamikus.”²⁵ A cikk – mint a példák is mutatják – egyedül a szó- és szófaji statisztikákkal foglalkozik Cselényi László és Juhász Ferenc versei kapcsán. De nemcsak az eredmények értékelése lehet kérdéses. Zsilka a szókincsgazdaság mérésére Pierre Guiraud²⁶ nemzetközileg elfogadott képletét alkalmazza:

¹⁹ Richard S. Forsyth, „Pops & Flops: Some Properties of Famous English Poems,” *Empirical Studies of the Arts* 18, 1. sz. (2000): 49–67.

²⁰ Justine Kao and Dan Jurafsky, „A Computational Analysis of Style, Affect, and Imagery in Contemporary Poetry,” *Linguistic Issues in Language Technology* 12, 3. sz. (2015), <https://nlp.stanford.edu/pubs/kaojurafsky12.pdf>. Lásd még Michael Dalvean, „Ranking Contemporary American Poems,” *Digital Scholarship in the Humanities* 30, 1. sz. (2015): 6–19.

²¹ Fekete István és Babarczy Anna, *Főnévi fogalmak konkrétsági, elképzelhetőségi és definiálhatósági értékeinek összefüggései*, kézirat, 2008.

²² Hász-Fehér Katalin, „A dilettantizmus kérdése a 19. század közepének kritikáiban,” *Acta Historiae Litterarum Hungaricarum* 32 (2016): 77–116.

²³ Karol R. Opara, „Grammatical Rhymes in Polish Poetry: A Quantitative Analysis,” *Digital Scholarship in the Humanities* 30, 4. sz. (2015): 589–598.

²⁴ Zsilka Tibor, *Stilisztika és statisztika* (Budapest: Akadémiai Kiadó, 1974), 19.

²⁵ Zsilka, *Stilisztika és statisztika*, 16.

²⁶ Guiraud munkájáról részletes ismertető olvasható magyarul: J. Soltész Katalin, „Guiraud statisztikai módszere a szókincs vizsgálatában” in *Általános nyelvészeti tanulmányok I.*, szerk. Telegdi Zsigmond (Budapest: Akadémiai Kiadó, 1963), 263–272.

$$R = \frac{V}{\sqrt{N}}$$

Egy másik tanulmányában ugyan jelzi, hogy Guiraud e képlettel végzett szókinccsgazdagsági vizsgálatait a francia nyelven írt szövegek esetében eredményeznek 20,5-ös átlagot,²⁷ de ezt az értéket mégis felhasználja Cselényi és Juhász költészetének összevetésére, mégis erre hivatkozva állítja, hogy a képlet nem használható 2000 szó alatti szövegek esetében, mert „a kapott eredmény nem egyezik az empirikus elvárás-szinttel,” azaz egy 742 szavas ciklus esetében a kapott 17,60 egy Kosztolányi-korpusztól kevés, mert „nem tükrözi híven a költő szókinccsgazdagságát.”²⁸ Ami éppenséggel lehet, hogy igaz, de mivel nem végzett kontrollvizsgálatot, nem nézte meg, 2000 szó fölött miként alakul ez az érték Kosztolányinál, valójában nem tudjuk, mekkora Kosztolányi verseinek TTR-értéke, a 17,60 valóban túl alacsony érték-e.

A kötet azonban még így is rendkívül inspiráló. A *Mérések a szöveg fonetikai, ritmikai és morfológiai szintjén* című tanulmánya 20. század elejének magyar költészetből vett példákat elemez.²⁹ A különböző képzési helyű magán- és mássalhangzók egymáshoz viszonyított arányainak, a szavak hosszúságának, szerkezeti fölépítésének vizsgálata ugyan öncélúnak tűnhet, de egyrészt valóban sikerül megragadni a stílusbeli különbségeket – másképpen: a különbségek magyarázhatók a stílusbeli eltérésekkel –, másrészt a háttérben Fónagy Iván nagy anyagot mozgató izgalmas vizsgálatai állnak, például a 19. századi francia költészet mássalhangzó-használata statisztikai és stilisztikai eredményeinek Petőfi költészetével való összevetése.³⁰ Ráadásul ez a szempont egyértelműen illeszkedik Berzsenyi versszemléletéhez is. A *Poétai harmonistika* egyik passzusa ugyanis így hangzik:

Harmóniába hozza a' poéta tárgyaival a' beszédet akkor, midőn stylusát tárgyai' természetéhez alkalmaztatja. [...] Kiterjed ezen harmóniázat még a' beszéd' külhangjaira is, úgy hogy a' poéta, valamennyire csak a' nyelv' természete enged, a' gyengébb érzelmekhez és szebb tárgyakhoz lágyabb és szebb hangu szavakat válogat, a' zordonabb tárgyakhoz pedig keményebb hanguakat.³¹

Azt sajnos nem lehet tudni, hogy Berzsenyi egészen pontosan mely hangokat gondolta lágyabbnak és szebbnek, melyeket pedig keményebbeknek. Fónagy Iván ugyan részletesen adatozza, hogy ez az elképzelés az ókortól kezdve a retorikai-poétikai gondolkodás része, de mivel nem teljesen egybeesnek az elképzelések, hiba volna bármelyiket is kiemelni. Egy levélrészlet ugyanakkor segít valamelyest közelebb lépni a kérdéshez. Berzsenyi egy Kazinczynak 1811 nyarán írt levélben helyesírási kérdésekről elmélkedve írja: „Annyi, mennyi, honnyi, olly, melly, st. mert ezeknek hangjok

²⁷ Zsilka, *Stilisztika és statisztika*, 16.

²⁸ Zsilka, *Stilisztika és statisztika*, 22.

²⁹ Zsilka, *Stilisztika és statisztika*, 46–75.

³⁰ Fónagy Iván, *A költői nyelv hangtanából* (Budapest: Akadémiai Kiadó, 1989), 36–42. Zsilka értelem-szerűen az első kiadást használta [1959].

³¹ Berzsenyi Dániel, *Prózai munkái*, kiad. Főrizs Gergely (Budapest: Editio Princeps, 2011), 388.

bizonytalan, vagy legalább a' különbség nem igen metsző a' kemény l, n, és lágy ly, ny között, és mivel a lágy ly többnyire még csak végső betűinkbe csuszott bé.”³²

Bármily változatos is a hangok jelentésével kapcsolatos hagyomány az ókortól a 19. századig (és tovább), az *l* és az *n* kemény volta nem szerepel benne – legalábbis Főnagyszavak anyagyűjtése alapján. Nehéz értelmezni Berzsenyi kijelentését. Egyedüli nyom, amely talán magyarázhatja az állítást, Révai stilisztikájának egyik szöveghelye. A „szép hanggal” kapcsolatos hibák („keménység,” „egyenlő hangozat,” „egyhangúság”) kapcsán olvashatjuk:

A' magános szavakban a sértő keménység minden nyelvekben a' mássalhangzóktól vagy on. A' mássalhangzók összeállító valóságos részeik ugyan a' szavaknak, mert tulajdonúl azokban áll minden jelentésök; a' magánhangzók pedig csak nyílásai a' száznak, mellyekel a' mássalhangzók kimondatnak.³³

Révai tehát azt állítja, hogy alapvetően minden mássalhangzó kemény, torlódásukat érdemes hát elkerülni, kivált, ha az adott mássalhangzók „már magokban is keményebbek” – teszi hozzá ugyanitt. Bár stilisztikájából nem derül ki, melyek ezek a „magokban” is keményebb mássalhangzók, Révai felfogása mégis magyarázhatja Berzsenyi kijelentését, és ami fontosabb, számunkra lehetőséget nyújt arra, hogy költészete sajátzerűségét megpróbáljuk megmérni – legalábbis egy vonatkozásban, a jólhangzás Révai ugyanezen passzusokban megfogalmazta követelményének és feltételeinek ismeretében. A tanulmány második részében erre tesz kísérletet, még ha ennek a hozadéka önmagában nem is túl sok, mivel egyelőre még hiányzik a háttér, azaz a korszak nyelvi jellemzőinek szélesebb, több szerzőt, több szövegtípust feldolgozó felmérése. Emellett a következőkben bemutatok még egy lehetséges elemzési módszert, amely egyes fogalmak felbukkanását, eloszlását és összefüggéseit kutatja. Ennek eredményei persze vita tárgyát képezhetik – egyáltalán, megfelelő fogalmakat vizsgáltam? –, de egyrészt jelzik a számítógépes vizsgálatok potenciálját, másrészt az eredmények elrendezésekor, vizualizálásakor felmerülő problémák is érdekes módszertani kérdéseket vetnek fel.

2. 1. A korpusz előkészítése

A vizsgálatokhoz a Magyar Elektronikus Könyvtárban található Berzsenyi-verseket választottam. Az Arcanum *Verstár* CD-je alapján készült verzió textológiai szempontból ugyan nem megfelelő, de mivel ez az interneten ma hozzáférhető magyar költészeti korpusz alapja, modernizálása egységes szempontok alapján és egyenletes minőségben készült, valamint az Országos Széchényi Könyvtár mint befogadó intézmény mégiscsak hitelesíti, végül úgy döntöttem, kiindulásként e kísérletben érdemes elfogadni.³⁴ Magam még annyit módosítottam a morfológiai elemzés pon-

³² Berzsenyi Dániel Kazinczy Ferencnek, Nikla, 1811. június 5. in Berzsenyi Dániel, *Levelezése*, kiad. Főrizs Gergely (Budapest: Editio Princeps, 2014), 222.

³³ Révai Miklós, *A magyar szép toll*, kiad. Éder Zoltán (Budapest: Akadémiai Kiadó, 1973), 52.

³⁴ A szövegek átírása mindig is a számítógépes vizsgálatok neuralgikus pontja volt. Lásd Vadai István kritikáját az első magyar számítógépes ritmuselemzésről: Vadai István, „Számítógép a verstan szolgálatában: Megjegyzések egy számítógépes ritmuselemzési kísérlet kapcsán,” *Irodalomtörténeti Közlemények* 88, 1. sz. (1984): 77.

tossága érdekében, hogy a kis- és nagybetűket a mondatstruktúrához igazítottam – kivéve értelemszerűen a tulajdonneveket –, valamint a versek tördelését is prózaivá alakítottam, mert a szövegek MEK-es kódolása sajnos üres sorokat tett be a versszakok közé, ami a morfológiai elemző számára megtévesztő volt. A legsúlyosabb, feltehetően a karakterfelismerés során elkövetett tévesztéseket javítottam (például *m* helyett *rn* – és fordítva). A verseket az 1816-os kötet sorrendjében mentettem le és neveztem el, azaz a vizsgálatok során a kötetrend dominált, az időrend ugyanis kevés kivételtől eltekintve hipotetikus, nagy intervallumokkal dolgozik, így informatív értéke meglehetősen csekély. Érdekes kérdés persze, hogy a számítógépes elemzések eredményét fel lehet-e használni ennek pontosítására, de ez most nem volt célom. A fogalmak korrelációinak vizsgálata a kötetrend fényében ugyanakkor hozzájárul a kötetkompozíció vizsgálatának kérdéséhez, új érveket szolgáltatathat az egyes elképzelésekhez, illetve segíthet megtalálni magukat a kulcsfogalmakat is.

A fogalomkeresés előtt a korpuszt a *Magyarlanc* elnevezésű nyelvi elemzővel preparáltam.³⁵ Az első oszlop az eredeti szöveget tartalmazza, a második a szóalakok lemmatizált változatát, a harmadik szófaját, a negyedik pedig a szóalak morfológiai elemzését (részletesebb, mondatszintű elemzéseket is képes a program készíteni). Mint a mellékelt képből látszik, a program alapvetően elboldogul a 19. század eleji versszöveggel is, még ha nem is sikerül mindig mindent pontosan megállapítania – erre később lesz egy konkrét példa is –, hiszen például a *legnemesebb* lemmája voltaképp a *nemes* kellene, hogy legyen, de azon a szinten, amire most nekünk kell, megfelelő.

2. 2. Fogalmak és összefüggéseik

A fogalom- vagy inkább kifejezéskeresést azért gondoltam Berzsenyi esetében értelmes lekérdezésnek, mivel – mint jeleztem – a kötet verseinek elrendezésével kapcsolatos értelmezésekhez szolgálhat adatokkal. Berzsenyi kötetkompozíciójával kapcsolatban Onder Csaba elképzelését vettem alapul.³⁶ Eszerint a kötet versei a „privatum” és a „publicum” szférájáról szólnak. A privatum meghatározó élménye a szerelem és a költészet, a publicumé pedig egy virágzó közösséggel szembeállított romló, hanyatló közösség. Ennek megfelelően három kulcsfogalom köré szerveztem a lekérdezéseket: az első a magyar/hon/haza szócsoporthoz (publicum), a második a múzsa/cüpris/camoena/szerelem (privatum), a harmadik pedig a tagadást kifejező nem/sem/ne/se/nincs. Ezek a fogalmak ugyan megfelelőnek tűntek, de huszonegy vers kimaradt, ti. sem az első, sem a második fogalomcsoport szavai nem fordultak elő bennük – feltételezve természetesen, hogy a versek vagy a privatum, vagy a publicum körébe tartoznak, harmadik csoport nincs. A következő versekről van szó: *A Melancholia*, *Jámborság és középszer*, *Gróf Török Sophiehez*, *A közelítő tél*, *Osztályrészem*, *Egy hívtelenhez*, *Chloe*, *Egy szilaj leánykához*, *Amathus*, *A csermelyhez*, *Egy leánykához*, *A megelégedés*, *Keszthely*, *Fohászkodás*, *Az est*, *Szerelmes bánkódás*, *Lilihez*, *Cencimhez*,

³⁵ János Zsibrita, Veronika Vincze and Richárd Farkas, „magyarlanc: A Toolkit for Morphological and Dependency Parsing of Hungarian,” *Proceedings of RANLP 2013* (2013): 763–771. *Magyarlanc*, hozzáférés: 2017.02.03, <http://rgai.inf.u-szeged.hu/index.php?lang=en&page=magyarlanc>.

³⁶ Berzsenyi Dániel, *Verseik: teljes gondozott szövegek: 1816*, szerk. Onder Csaba, Matúra Klasszikusok (Budapest: RaabeKlett, 1998), 21–37.

95	Boldog	boldog	ADJ	Case=Nom Degree=Pos Number=Sing	
96	vagy	van	VERB	Definite=Ind Mood=Ind Number=Sing Person=2 Tense=Pres VerbForm=Fin Voice=Act	
97	,	,	PUNCT	–	
98	Áon	Áon	PROPN	Case=Nom Number=Sing	
99	szüzei	szüz	NOUN	Case=Nom Number=Plur Number[psor]=Sing Person[psor]=3	
100	kedvese	kedves	NOUN	Case=Nom Number=Sing Number[psor]=Sing Person[psor]=3	
101	,	,	PUNCT	–	
102	S	s	CONJ	–	
103	nagy	nagy	ADJ	Case=Nom Degree=Pos Number=Sing	
104	,	,	PUNCT	–	
105	mint	mint	SCONJ	–	
106	hazádnak	haza	NOUN	Case=Dat Number=Sing Number[psor]=Sing Person[psor]=2	
107	legnemesebb	nemesebb	ADJ	Case=Nom Degree=Sup Number=Sing	
108	fia	fia	NOUN	Case=Nom Number=Sing Number[psor]=Sing Person[psor]=3	
109	!	!	PUNCT	–	
110					
111	E	e	PRON	Case=Nom Number=Sing Person=3 PronType=Dem	
112	két	két	NUM	Case=Nom NumType=Card Number=Sing	
113	remek	remek	ADJ	Case=Nom Degree=Pos Number=Sing	
114	dísz	dísz	NOUN	Case=Nom Number=Sing	
115	kéri	kér	VERB	Definite=Def Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin Voice=Act	
116	méltán	méltán	ADV	–	
117	A	a	DET	Definite=Def PronType=Art	
118	Ganyméd	Ganyméd	PROPN	Case=Nom Number=Sing	
119	poharát	pohár	NOUN	Case=Acc Number=Sing Number[psor]=Sing Person[psor]=3	
120	az	az	DET	Definite=Def PronType=Art	
121	égben	ég	NOUN	Case=Ine Number=Sing	
122	.	.	PUNCT	–	

1. ábra. A korpusz elemzése a *Magyarlanc* alkalmazásával

Esztihez, Báró Wesselényi Miklós képe, A temető. – Itt valójában indulhattam volna másik irányba is – hiszen az ilyen fogalomkeresés nem más, mint számítógéppel megsegített *close reading* –, ti. ráhagyatkozom a számítógépre, és megnézem, miként csoportosítja maga a verseket, majd ebből megpróbálok következtetéseket levonni. A dolgozat egy későbbi pontján ezt röviden bemutatom.

A hiba egyik oka, hogy a *Magyarlanc* nem tudta helyesen elemezni a *múzsát*, *múzsához* szóalakokat és egy *múzs* alapszót tételezett. A mai magyar nyelvre épülő, meglehetősen nagy, 1200000 szót és 250000 írásjelet tartalmazó, hat területről (szépirodalom, jog, hír- és hetilapok cikkei, gazdasági hírek, kamaszoktól származó szövegek, informatikai irodalom) összegyűjtött, nyelvészek által morfológiailag elemzett tanító korpuszában nyilván nem szerepelt ez a kifejezés.³⁷ Tehát ezt módosítani kellett – persze ha alanyesetben szerepel a versben, akkor az elemző is *múzs*a-ként azonosítja. Ezt a csoportot kiegészítettem még az *ifjú* és *kebel* kifejezéssel, illetve a *cüpris*, *camoena* nagybetűs alakját is beemeltem a biztonság kedvéért, mert a *Magyarlanc* nem minden szót kisbetűsített. Az *isten* viszont például nem volna elég specifikus, mivel nagyon gyakran fordul elő az első csoport verseiben is (például „honnunk isteni”). Hat vers viszont még mindig kihullott. Ezért felvettem ide még a *mirtus*-t, az első csoportba pedig a *bárá*-t és a *Rómá*-t, de négy vers, *A közelítő tél*, *Egy hívtelenhez*, *A csermelyhez*, *Fohászkodás* ezek után is fennmaradt. A második csoportba tettem ezek után a *fohász*-

³⁷ Csak a Szeged Dependency Treebank adatait ismertettem, jöllehet a *Magyarlanc* tanító készletében ezenkívül még a szintén kézzel annotált Szeged Corpus is szerepel. Zsibrita, Vincze and Farkas, „magyarlanc,” 763–771.

kodás-t, valamint a *labirintus*-t is, ezzel sikerült minden verset valamelyik csoportba osztani.

Mint jeleztem, a fogalmak megfeleltetése kérdés tárgya, de kiinduló elemzésnek, mintának, megfelelnek. Az érdekelt, van-e valami kimutatható, számszerűsíthető összefüggés közöttük, illetve mutat-e különbséget az eredmény ahhoz képest, mintha véletlenszerűen rendeznénk el a szövegeket, azaz a versekhez véletlenszerűen rendelnénk, hogy a privatum vagy a publicum szavai fordulnak elő bennük. A vizsgálatot értelemszerűen a lemmatizált alakokon végeztem a statisztikai elemzések, illetve a természetesnyelv-feldolgozás során gyakran használt R programozási nyelven, és mindig az R beépített függvényeit, parancsait használtam az értékek kiszámításához.³⁸

Mint a mellékletben látható, ha a nyers szöveg versenkénti bontásban rendelkezésünkre áll, elég kevés, mindössze huszonegy parancssorral megoldható a vizsgálat – persze a lényeg, az értelmezés, ezután következik. Az első táblázat egy részlet, amely azt mutatja, a kötet első öt szövegében a keresett kifejezések előfordulnak-e vagy sem (az első esetet „1”, a másodikat „0” jelöli). A táblázat természetesen mind a nyolcvanhét versre vonatkozó adatokat tartalmazza: az első szócsoporthoz összesen 31, a második 73, a harmadik 66 versben van jelen, tehát lesznek olyanok, amelyekben mind a privatum, mind a publicum szavai előfordulnak. Érdekesebb azonban a második táblázat, amely azt vizsgálja, van-e összefüggés a fogalmak felbukkanása között (az értékeket az átláthatóbb eredmények érdekében négy tizedesjegyig adom meg, az R tíz tizedesig számítja ki).

	magyarok	istennők	nemek
<i>Ajánlás</i>	1	1	0
<i>Küprishez</i>	0	1	0
<i>A melancholia</i>	0	1	0
<i>A szerelemhez</i>	0	1	1
<i>A jámborság és középszer</i>	0	1	1

	magyarok	istennők	nemek
magyarok	1,0000	-0,5886	0,1393
istennők	-0,5886	1,0000	-0,0277
nemek	0,1393	-0,0277	1,0000

1. táblázat. A vizsgált fogalmak előfordulása

De mit jelent mindez? A korreláció értéke -1 és +1 között lehet. Ha az előbbi, akkor teljes a korreláció, de fordított, tehát ha az egyik érték nő, akkor a másik ugyanolyan mértékben csökken, ha az utóbbi, akkor mindig ugyanabba az irányba, ugyanolyan mértékben változnak. Minél közelebb van egy érték a nullához, annál kisebb a korreláció, nulla esetében egyáltalán nem lehet kimutatni a két fogalom előfordulása

³⁸ A vizsgálat megismételhetősége/ellenőrizhetősége érdekében a használt kódok kommentárokkal kísért listája az interneten is olvasható a <http://labadigergely.github.io/szovegek/2017/02/17/R5/> címen. Az értékek kiszámításakor az R által használt képletek a program súgójából megismerhetők. R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing (Vienna, 2016), hozzáférés: 2018.05.23, <https://www.R-project.org/>.

között bármiféle összefüggést. A ± 0.5 és a ± 1 közötti érték esetében kifejezetten erős korrelációról szoktunk beszélni, a ± 0.1 és ± 0.3 közöttiekénél van ugyan, de gyenge a korreláció – a Pearson-féle korrelációs koefficiens értelmezésében Jockers könyvét követem.³⁹

A vizsgált fogalmak esetében a „magyarok” címszóval összefoglaltak és az „istennők” között ténylegesen kimutatható, és kifejezetten erős negatív kapcsolat van ($-0,5886$), tehát azokban a versekben, amelyekben az első csoport szavai/témái előfordulnak, a másodiké nem – és fordítva: a privatum és a publicum világa nem keveredik. Szintén kimutatható a kapcsolat a „magyarok” és a „nemek” között ($0,1393$), azaz a publicum világában a tagadás, a hiány jelen van, de ez az érték alacsony, a korreláció gyenge. A „nemek” és az „istennők” között viszont gyakorlatilag nincs semmiféle kapcsolat ($-0,0277$), azaz a privatumban a hiány nem meghatározó élmény – persze a keresett kifejezésektől, ismétlem, sok függ.⁴⁰

Mindez így önmagában nem több, mint Onder kötetkoncepciójának számszerű igazolása, ami éppenséggel nem kevés, de tény, nyolcvanhét vers nem akkora adatmennyiség, amekkorát egy ember ne tudna feldolgozni. Innen több irányba is tovább lehet lépni: más fogalmakra keresve talán más koncepciót erősebben lehet igazolni/cáfolni. De talán izgalmasabb, látványosabb, ha az eredményt randomizáltan ellenőrizzük, azaz megnézzük, hogyha az egyes versekben előforduló kifejezések értékeit összekeverjük – tehát az első táblázat nulláit és egyeseit véletlenszerűen osztjuk ki a versek között –, miként alakulnak a korrelációs értékek, mert ekkor derül ki, hogy a tényleges eredmény mennyire egyedi, váratlan, vagy ha úgy tetszik, a szerző mennyire volt „tudatos.” Az R által kínált lehetőségekkel élve természetesen nem nekünk kell újra és újra kiosztanunk a számokat, ezt elvégzi a gép. Miután a program tízezerszer véletlenszerűen kiosztotta, hogy melyik vers mely csoportba tartozzon, a következő eredményeket kapjuk az összefüggésekről – mivel korrelációt csak két érték között tudunk számolni ezért a következő diagram csak a publicum és a privatum lehetséges korrelációit mutatja be, de mint az utána következő táblázatból kiderül, természetesen mindegyik párosra kiszámoltattam az értékeket.

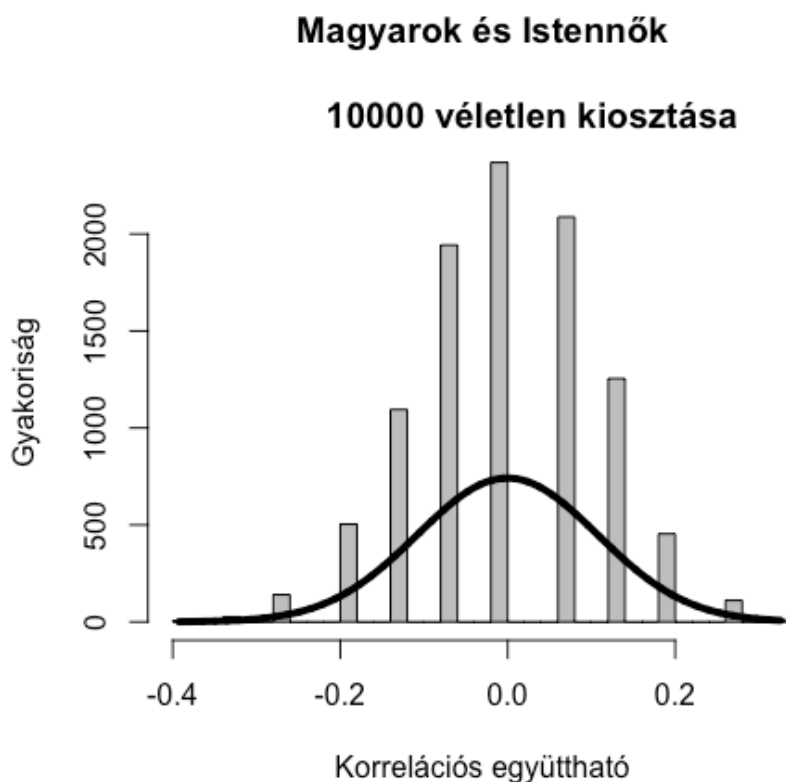
Bár a lényeg talán ebből is látszik, az alapvető statisztikai adatok kiírása egyértelműbbé, és könnyebben elemezhetővé teszi az eredményeket, ráadásul mindet tartalmazza:

	min.	max.	átlag	szórás	tényleges érték
magyarok–istennők	-0,3926	0,3258	-0,0002	0,1076	-0,5886
magyarok–nemek	-0,3655	0,3636	-0,0002	0,1077	0,1393
nemek–istennők	-0,2470	0,4840	0,0003	0,1083	-0,0277

2. táblázat. A korrelációs értékek számszerűsítve páronkénti bontásban

³⁹ Matthew L. Jockers, *Text Analysis with R for Students of Literature* (New York: Springer, 2014), <https://doi.org/10.1007/978-3-319-03164-4>, 49. A korrelációanalízis ötletét innen vettem, viszont a kódok jelentős részét én írtam, lévén Jockers más jellegű korpuszon és más módszertannal dolgozott: egy regény, a *Moby Dick* fejezeteit és két kifejezésének (*whale*, *ahab*) darabszám szerinti és százalékos előfordulását és összefüggését vizsgálta.

⁴⁰ Lehetséges, hogy ebben a kérdésben a számítógépes nyelvészeti használatos szentiment- és emócióelemzés részletesebb válaszokat tudna adni.



2. ábra. A publicum és a privatum korrelációi

A keresett szavak előfordulásának tízezer véletlenszerű kiosztása esetén az átlagos korreláció minden esetben lényegében nulla, azaz nincs semmiféle összefüggés a szavak versekbeli előfordulása között. A szórás is nagyon alacsony, éppen eléri a 0,1-es határt, amelytől valamiféle gyenge kapcsolatról már lehet beszélni, mindez azt jelenti, hogy a tízezer eset túlnyomó részében az adatok nulla körül maradnak. Ehhez kapcsolódóan van egy statisztikai szabály is, a 68–95–99,7, amely azt jelzi, hogy a mért adatok 68%-a az átlagtól legfeljebb egy szórásnyira tér el, a 95%-uk az átlagtól kétszórásnyi távolságon belül helyezkedik el, 99,7%-uk pedig legfeljebb háromszórásnyi távolságban található az átlaghoz képest, így a szórásérték a tényleges adat váratlanságát, valószínűségét mutatja. A szélsőértékek persze azt is megmutatják, hogy ilyen alacsony esetszám (31, 73, 66) és kevés érték (0, 1) esetén nehéz olyan elosztást találni, amelynél igazán erős korrelációról beszélhetnénk. Igaz, a fenti táblázat szélsőértékei majd mindegyik esetben kívül esnek a szélsőértékek háromszorosán (egyedül a nemek–istennők minimumértéke van belül a háromszorosán), azaz tízezer eset háromtized százaléka (30), ha tényleges eredményként jelenne meg, váratlannak volna minősíthető. Éppen ezért fontos, hogy a „magyarok” és az „istennők” közötti tényleges adat még a tízezer randomizált kiosztás minimumértékén is jóval túl található, tehát a statisztikai előrejelzés szerint teljességgel váratlan és valószínűtlen a ténylegesen mért adat. Ami az egyébként már önmagában is erősként interpretálható korrelációt (-0,5886) még erősebbé teszi, azaz a privatum és a publicum – az általam használt fogalmakkal mért – koncepciója erős támogatást kap, hiszen a tényleges értéket

tízezer próbálkozásból egyszer sem sikerült előállítania a programnak. A „magyarok” és „nemek” között mért kapcsolat távol van a lehetséges szélsőértékektől, de belül a szórás kétszeresén, ami a körülmények ismeretében (alacsony érték- és esetszám) a mutatótnál (0,1393) talán valamivel erősebb összefüggést valószínűsít. A „nemek” és az „istennők” közötti -0,0277-es érték már eleve a kapcsolat teljes hiányát mutatta, amit a randomizálás is megerősített: bár a mért adat magasabb, mint az átlag, ám jócskán belül van az egyszeres szóráson. Figyelembe véve azonban, hogy a szélsőértékek itt is ugyanakkora intervallumot fognak át, mint a másik két esetben, tehát akár erősebb összefüggés is elképzelhető volna, a kapcsolat teljes hiányát igazoló tényleges adat megint a koncepciót erősíti – vagy a csoportosítás során jól megválasztott fogalmakat dicséri.

Természetesen további vizsgálatokat is el lehetne végezni. Ha magukon a fogalmakon nem is változtatunk, még akkor is értelmesnek tűnik az egyes könyveken belül kiszámolni az értékeket vagy visszalépni az időben és Berzsenyi 1808-as kötetkompozíciójával vetni össze az eredményeket – az induló kódkészlet most már mindenki rendelkezésére áll. A legkézenfekvőbb azonban, ha megnézzük, hogy a számítógépes algoritmusok milyen tematikus csoportokat találnak a nyolcvanhét versben, ezek igazolják-e a koncepciót. Egy módszertani megjegyzést kell azonban előre bocsátanom: a *topic modelling* hosszabb szövegek esetén megbízhatóbb. A kísérlet ugyanakkor érdekes eredményt adott még így is.

A tartalomelemzés során meg kell adni, hány témát keressen az R algoritmus – én a *tm* csomagot használtam⁴¹ –, értelemszerűen kettőt adtam meg, és nem a versekben ténylegesen előforduló szavak, hanem ezek szótári formái alapján keresett a program. Az első csoport első ötven szava: *szent, kéz, lát, égi, föld, néz, bölcs, dicső, isten, világ, nap, magyar, nép, ember, gyenge, vér, haza, lélek, örök, vad, boldog, virág, érez, harc, tesz, barát, víg, halál, ész, fény, róma, század, fej, fényes, kor, tenger, elme, kény, kevély, kincs, por, száll, tűz, zár, bér, gyönyörű, kebel, múzsa, pálya, szabad*. A sorrendet a gép állította fel, s minél előrébb van egy szó, annál jellemzőbb az adott témára. A második csoport első ötven szava: *szív, öröm, lélek, szem, bús, szerelem, édes, élet, kar, kép, megy, könny, szelíd, hív, láng, arany, idő, tündér, kedves, szeret, szűz, csendes, homály, kegyes, liget, mély, forró, jer, szerető, kebel, mosolyogva, tud, vár, zöld, magas, völgy, kér, mennyei, erő, gond, hall, int, lel, mosolyog, rózsza, tér, vesz, vidám, fed*.

Ezzel a módszerrel le kellett mondanom arról, hogy a publicum világában (első csoport) jelen lévő hiányt vizsgáljam, mivel a tagadószavakat mint tipikus funkciószavakat a *topic modelling* nem veszi figyelembe. A két listát összevetve látszik, hogy az eredeti koncepció publicum–privatum megoszlása érvényes, de a tartalmuk – legalábbis az általam fentebb adott összefoglalóhoz képest – megváltozott. A múzsaverseket, a tudomány dicséretét tartalmazó darabokat a publicum körébe osztja az algoritmus – ami abból a szempontból aligha meglepő, hogy a megénekelt kortárs személyiségek érdemei közt az ősi hadi erények mellett a művelődésben betöltött szerepüket is kiemelik az ódák –, a privátumban pedig a szerelmi költészet mellett hangsúlyosan jelennek meg a sztoikus életfilozófiát, életvilágot bemutató versek. Ha mindezt az egyes darabokra lebontva kívánjuk vizsgálni, akkor jóval árnyaltabb képet kapunk,

⁴¹ Ingo Feinerer, Kurt Hornik and David Meyer, „Text Mining Infrastructure in R,” *Journal of Statistical Software* 25, 5. sz. (2008): 1–54, <http://www.jstatsoft.org/v25/i05/>.

mint a fogalomkeresős módszer esetében, hiszen a *topic modelling* kiindulópontja, hogy minden szöveg több témából áll. Azaz valamilyen mértékben minden versben megtalálható mindkét téma. Nézzük meg, hogy alakulnak az előző táblázat értékei a tartalomelemzés során:

	publicum	privatum
<i>Ajánlás</i>	59,67%	40,32%
<i>Küpriszhez</i>	42,67%	57,33%
<i>A melancholia</i>	37,78%	62,22%
<i>A szerelemhez</i>	38,46%	61,54%
<i>A jámborság és középszer</i>	49,65%	50,35%

3. táblázat. Az értékek alakulása a tartalomelemzés során

Az *Ajánlás*-ban tehát az első téma dominál közel 60%-kal, a többi négyben viszont a második. Maguk az adatok tehát alátámasztják az előző táblázat 0-it és 1-eit, ha 50%-nál húzzuk meg a határt. A *jámborság és középszer* értékei azonban a gépi olvasás sajátos mechanizmusára is felhívják a figyelmet. Egy értő olvasó aligha sorolná (kerekítve) 50%-ban a publicum világába tartozónak a szöveget, a program számára azonban a kulcsszavak megléte a kérdés – a „szent,” a „templom,” az „áldozópap,” a „rabiga” valóban értelmesek a közösségi lét diskurzusának keretében. A vers értékei mindemellett talán Berzsenyi alkotómódszerére nézve is érdekesek, mennyire könnyen vándorolnak egyes képzetek, képek a különböző időpontban, műfajban és céllal született versek között. E vers mellett a *Küpriszhez* értékei talán még meglepőbbek: nehéz értő szemmel 43%-ban a publicum világába sorolni – persze például a „bér,” a „tanít” megint csak részesei a másik diskurzusnak, tehát nincs szó tévedésről. Inkább az a kérdés ezek után, hol húzzuk meg a határt, azaz hány százaléknál minősítsünk egy szöveget egyértelműen az egyik vagy a másik csoportba tartozónak. Ha 50%-nál, akkor az első témába 28, a másodikba 59 szöveg tartozik; 60% esetében 13, illetve 74. A publicum körébe tartozó szövegek közül legnagyobb értékkel *A pesti magyar társasághoz* rendelkezik, a számítógépes algoritmus szerint 75,47%-ban tartozik ide, s csak 24,53%-ban a privátumba. Ebben legnagyobb értékkel a *Lilihez* („Jer, Lili! nézd...”) rendelkezik (78,16%).

2. 3. Magán- és mássalhangzók, szóhosszúság

A tanulmány befejező részében a „szép hanggal” kapcsolatos stilisztikai kérdésekre próbálok meg válaszokat adni. A korábban idézett Berzsenyi-passzus és a vele összekapcsolt Révai-részlet nyomán többféle szempont alapján is lehet a verseket vizsgálni. Az egyik a magán- és mássalhangzók aránya, mivel ezek „illő öszvemérsékeltetése szerzi főképen minden nyelvnek az ő kedves hangját.”⁴² A másik szintén könnyen mérhető tulajdonság a szavak hosszúsága, amely túlzott monotonitás esetén szintén „keménység”-et okoz Révai szerint.

A szövegeket értelemszerűen most a tényleges szóalakok alapján kell vizsgálni, ami még élesebben veti fel a szöveg-előkészítés nehezen megválaszolható módszertani kérdését, a helyesírás normalizálását, és óhatatlanul lesznek hibák. A program

⁴² Révai, *A magyar szép*, 52.

alapvetően csak a latin betűket ismeri fel, be lehet ugyan állítani, hogy például a *th*-t, *ph*-t, vagy éppen az *ae*-t egy hangzóként értse, de az ilyen megoldások nem mindig elegendők. A *vaszár* ebben a korpuszban szerencsére nem fordul elő (csak acélból készült záruk, závarok), de egy szimpla gépi számolásnál a helyes mássalhangzószám a *vaszár* esetében négy volna, viszont ha beállítjuk az *sz*-et, akkor a program hármat fog találni. A korban ugyan az *s*-nek volt egy másik karaktere is (*/*), amelyet Verseghy szerint arra kell használni, hogy az összetett mássalhangzókban (*sz*, *zs*) az *s*-et jelölve a szóösszetételekből adódó félreértéseket el lehessen kerülni, illetve különböző szabályokat lehessen tanítani a diákoknak kiejtésről, elválasztásról,⁴³ ám ezt az átiratok nem őrzik meg – és az egykorúak közül sem mindenki, nem minden szerző, kiadó, nyomda alkalmazta ennyire következetesen. A hosszú összetett mássalhangzókat viszont nem kódoltam külön, mert az értékük mindenképpen kettő lesz: a gép persze azért számolja annak az *lly* esetében, mert az *l*-et és az *ly*-et észleli, az ember pedig azért, mert tudja, hogy az *ly* duplázódik. Ezzel természetesen le kell mondani az egyes mássalhangzók ismétlődésének vizsgálatáról – legalábbis nekem. És akkor még nem beszéltünk arról, hogy az *y*-t hová számoljuk: bizonyos mássalhangzók mellett *ü*-ként vagy *i*-ként kell számolni, mások mellett pedig egy összetett mássalhangzó része. Ezt viszont legalább meg lehet oldani egy erre is ügyelő átirással. Mindezt szükséges még kiegészíteni azzal Jékel és Papp munkája nyomán,⁴⁴ hogy bár a „szép hanggal” kapcsolatos vizsgálódást emlegettem korábban, valójában fonémákat fogok vizsgálni, hiszen az írásban nem jelölt összeolvadásokat, részleges és teljes hasonulásokat nem vettem figyelembe – a különbség Jékelék ellenpróbái szerint elhanyagolható, mindössze 4 ezrelék⁴⁵ – a címetek azonban velük ellentétben igen.

A számolást ismét hamar elvégzi a gép, eredményét az alábbi táblázatban foglaltam össze. Az érdekesebb rész, mint mindig, utána következik.

	Ajánlás	Első könyv	Második könyv	Harmadik könyv	Negyedik könyv
msh	354	8296	12588	12707	9520
mgh	250	5643	8548	8686	6341
átlag szóhossz	5,2983	5,1722	5,1189	5,1711	4,8713
msh/mgh	1,4160	1,4701	1,4726	1,4629	1,5013
min. msh	0	0	0	0	0
max. msh	8	9	10	10	10
átlag msh	3,1053	3,0783	3,0487	3,0716	2,9238
szórás msh	1,1708	1,6408	1,6805	1,7120	1,7132
min mgh	0	0	0	0	0
max mgh	5	6	6	7	6
átlag mgh	2,1930	2,0939	2,0702	2,0996	1,9475
szórás mgh	1,1279	1,1046	1,1021	1,1259	1,1041

4. táblázat. A fonémavizsgálat eredménye

⁴³ Verseghy Ferenc, *Magyar grammatika avagy Nyelvtudomány* (Buda: Egyetemi Nyomda, 1818), 56.

⁴⁴ Jékel Pál és Papp Ferenc, *Ady Endre összes költői műveinek fonémastatisztikája* (Budapest: Akadémiai Kiadó, 1974).

⁴⁵ Jékel és Papp, *Ady Endre összes*, x.

A táblázat sorai közül talán egyedül az átlagos szóhossz szorol magyarázatra: összeadja az adott szakasz mással- és magánhangzóinak számát, majd elosztja a szavak számával. A szavak átlagos hosszai nagyon közel esnek egymáshoz, ami a mással- és magánhangzók aránya kiegyenlítetttségének fényében persze nem meglepő. Kérdés, ez mennyire az egyéni nyelvhasználat jellemzője, vagy mennyire befolyásolhatja maga a mérés. Az utóbbit kizárhatjuk, mert a keresés során harminc mássalhangzót adtam meg és csak 16 magánhangzót,⁴⁶ amelyek aránya 1,875, ami jóval magasabb, mint a mért 1,47-es, 1,5-ös érték. Mielőtt kontrollanyaggal vetném össze, a táblázat értelmezéséhez érdemes még hozzátenni, hogy az egyes hangzócsoporthoz közelebb esik az átlaga a legalacsonyabb szélsőértékhez, és az ezzel párosuló magas szórás érték azt jelzi, hogy kevés a kifejezetten rövid szó – amit Révai expliciten tilt is –, viszont ha megnézzük, a szórás háromszoros értéke közel van a maximumhoz, de sosem éri el. Azaz a szavak meglehetősen változatosságot mutatnak a szóhosszúságot tekintve, és az extrémhosszú, tíz mássalhangzót vagy hét magánhangzót tartalmazó szavak, legfeljebb a szöveg 0,3%-át teszik ki – a túlzottan hosszú szavaktól is óv Révai. Tehát a versek a Révai-féle jólhangzás követelményeinek (változatos szóhossz, magán- és mássalhangzók arányos eloszlása) megfelelnek.

Kontrollként megvizsgálva egy Berzsenyi-levét, valamint -értekezésrészlet, illetve néhány Virág-vers és -értekezésrészlet szövegét,⁴⁷ a következő táblázatot kapjuk:

	Berzsenyi- értekezés	Berzsenyi- levél	Virág- értekezés	Virág-versek
msh	1435	1229	2904	994
mgh	961	2025	2025	671
átlag szóhossz	5,3363	5,6590	5,6590	4,9554
msh/mgh	1,4932	1,4341	1,4341	1,4814
min. msh	0	0	0	0
max. msh	9	10	10	9
átlag msh	3,1960	3,1035	3,3341	2,9583
szórás msh	1,8915	1,8751	2,0633	1,7151
min mgh	0	0	0	0
max mgh	6	6	6	5
átlag mgh	2,1403	2,1035	2,3250	1,9970
szórás mgh	1,1497	1,2657	1,2789	1,0968

5. táblázat. Berzsenyi- és Virág-szövegek fonémavizsgálatának eredménye

Szembetűnő, hogy a vizsgált Virág-versek adatai egy kivétellel teljesen összhangban vannak a Berzsenyi-kötet adataival. A magánhangzók szórásának háromszoros értéke

⁴⁶ A hangzók száma a *th*-, *ae*-típusú kiegészítések miatt nőtt meg. A függelékben olvasható kódban valamennyi látszik.

⁴⁷ A levél kivételével MEK-en elérhető szöveget vettem alapul. Az értekezésrészletek: Berzsenyi Dániel: *A magyarországi mezei szorgalom némely akadályairól*; Virág Benedek: *Magyar századok*. A Virág-versek a következők: *Gróf Forgács Miklós emlékezete*, *Bosszúállás*, *Egy jeles poétára*. A levél pedig Berzsenyi fent idézett levele Kazinczynak 1811 nyaráról a kritikai kiadás alapján. A feldolgozás a versekkel megegyező módon történt.

az átlaghoz adva eléri a maximumértéket, míg a Berzsenyi-verseknél ez egyik esetben sem történt meg. Azaz a Berzsenyi-versek szóhosszai valamivel változatosabbak. Virág prózai szövegeinek magánhangzósórása szintén ugyanezt mutatja (a háromszoros értéke az átlaghoz hozzáadva meghaladja a magánhangzók szélsőértékét), de Berzsenyié már nem. Érdekes továbbá, hogy a prózai szövegek esetében az átlagos szóhossz valamivel nagyobb, mint a versek esetében, Virág szövegénél ez elég markáns különbség, a másik két esetben legfeljebb tendenciáról beszélhetünk – elképzelhető, hogy a témaválasztás volt hatással erre az értékre, ezt a későbbiekben érdemes megvizsgálni. A mással- és magánhangzók arányában azonban nem látszik különbség. Egy nagyobb kontrollanyag alapján a konklúzió is erőteljesebb lehetne, de mindezek után úgy tűnik, Berzsenyi a jóhangzás követelményének jegyében mind prózában, mind versben odafigyelt a szóhosszúság változatosságára, a mással- és magánhangzók egyenletes eloszlása azonban sokkal inkább nyelvi sajátságnak tűnik, semmint a szerzői nyelvhasználat jellemzőjének.

3. Következtetések

A tanulmányból kiderült, hogy a gépi olvasás mechanikus voltának megvannak a maga gyenge pontjai – emlékezzünk például az *lly* értékének kérdésére –, de még egy ilyen kis korpusz esetén is látszik a potenciálja. A versek gépi, tematikus csoportosítása után felmerült, hogy a vizsgált koncepció tartalmát esetleg érdemes módosítani, hogy a művelődés problémái, illetve a sztoikus világlátás kérdései hangsúlyosabban jelenjenek meg a Berzsenyi-versek értelmezésében. De mindkét vizsgálat egyértelművé tette: 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 biztonságosan használni.

The Reading Machine

A Distant Reading of Dániel Berzsenyi's Poems

This experimental study is an attempt to explore and test whether, by using current computational linguistics tools for the Hungarian language, it is possible to confirm hypotheses concerning the classification of Dániel Berzsenyi's (1776–1836) poems drawing on their close reading. The paper also investigates if we can answer questions related to the phoneme theory and practice in Berzsenyi's poems, which obviously cannot be examined by traditional close reading methods.

Keywords:

Dániel Berzsenyi, R, Magyarlanc, topic modelling, distant reading

Drótos László

Országos Széchényi Könyvtár

mekdl@iif.hu

Kokas Károly

SZTE Klebelsberg Könyvtár

kokas@ek.szte.hu

Webarchiválás és a történeti kutatások

A digitálisan születő tartalom sokkal részletesebb és teljesebb leképezése a jelennek, mint ami régebbi korokban a hagyományos információhordozó eszközökkel rögzíthető volt. A tanulmány első része arról ad áttekintést, hogy milyen próbálkozások és technológiák léteznek ennek a digitális jelennek a megőrzésére, illetve milyen korlátai vannak a már működő webarchívumoknak. A dolgozat második része azt vizsgálja, hogy a történeti szempontú kutatásoknak hogyan lehet hasznára mindez, s hogyan lesz elsősorban a közelmúlt történetének is elsőrangú forrása. A szerzők arra is rámutatnak, hogy a webaratások következtében előálló hatalmas adatsilók egészen új típusú forráskezelést és módszertant kívánnak majd meg, miközben azzal kecsegtetnek, hogy egészen új típusú eredményeket is fel lehet majd mutatni segítségükkel.

Kulcsszavak:

webarchiválás, digitális megőrzés, digitális bölcsészet, webhistoriográfia



A történelmet már nagyrészt online írják.¹

Olyan mértékben függünk tőle, mint az elektromos hálózattól, de amíg azon csak energia folyik, az interneten információ, mégpedig az élet minden területéről, az intim magánügyektől a globális közügyekig. A világhálón áramló elektromágneses impulzusok minden pillanatban minden korábbinál részletesebb lenyomatát adják civilizációnk jelenének. Ezeket a lenyomatokat valahogyan meg kellene őrizni ahhoz, hogy a múlttá váló jelenünk értelmezhető legyen a jövőből visszanezve. Két-három évtizedet már elvesztettünk... bár azért nem teljesen.

¹ „De geschiedenis van vandaag wordt vooral online geschreven.” Peter de Bode, René Voorburg, „Webarchivering,” hozzáférés: 2018.05.22, <https://www.kb.nl/organisatie/onderzoek-expertise/e-depot-duurzame-opslag/webarchivering>.

1. Ki őrzi meg a netet?

A digitális univerzum elképesztő táguási üteme elfedi azt a tényt, hogy ez a világ rendkívüli tempóban pusztul is. A tudományos publikációkban hivatkozott internetes források esetében öt és tíz év közötti felezési értéket mutattak ki a különböző vizsgálatok, vagyis ennyi idő alatt a linkek fele eltörik, eltűnik mögülük az eredeti tartalom. A *webkettes* helyeken még gyorsabb az erózió, egyes Facebook-posztok, Twitter-üzenetek, YouTube-videók élettartama csak napokban mérhető. A 404-es hibaüzenet a legtöbbször megtekintett internetes tartalom.²

Az 1970-es és 1980-as évtizedek hálózatainak: a BBS-eknek, a CompuServe-nek, a Minitelnek, az EARN/BITNET-nek, a nálunk is elterjedt X.25-nek, vagy a korai Internetnek a bitjei és bájtjai már nagyrészt eltűntek. Az 1980-ban indult Usenet fórumainak üzeneteit 1995-től elkezdte egy Deja News³ nevű vállalkozás gyűjteni és megőrizni, majd 2001-ben a Google megvette az akkorra már több mint félmilliárdnyi üzenetből álló gyűjteményt, kiegészítette más forrásokból 1981 májusáig visszamenőleg, és beolvasztotta a saját Google Groups szolgáltatásába. A kilencvenes évek első felének legnépszerűbb internetes tartalomszolgáltató felülete, a Gopher is majdnem teljesen eltűnt a sülyesztőben, ha 2007 júniusában egy John Goerzen nevű programozó le nem menti a még működő szervereket, és a körülbelül 780 ezer dokumentumot tartalmazó 15 gigabájtos tömörített csomagot oda nem adja az Internet Archive-nak.⁴ Sajnos ez egy megkésített akció volt, mert az internettörténeti szempontból legérdekesebb Gophereket akkorra már rég leállították.

A web esetében szerencsére kisebb volt a kérés. Öt évvel az első weboldal megszületése után, 1996-ban San Franciscóban létrejött az Internet Archive nevű nonprofit szervezet és archívum, amely az egyéb digitális média (könyv, kép, hang, videó, szoftver) mellett gyűjti a webhelyek tartalmát is. Ezeket részben a jelenleg már az Amazon cégcsoportjába tartozó és főként az internetes oldalak forgalmának mérésével és rangsorolásával foglalkozó Alexa Internet cégtől kapja. 2017 elején 279 milliárd weboldal volt a gyűjteményben, melyek különböző időpontokbeli mentései visszanezhetők a Wayback Machine⁵ nevű szolgáltatással. De ez a hatalmas szám is csak töredéke a teljes webtérnek, mert sok webhely bejárhatatlan és lementhetetlen automatikus módszerekkel: vagy mert olyan technológiát használ, vagy mert jelszóval védett, vagy csak egyszerűen ki vannak róla tiltva a robotok.

A közösségi média különösen nehezen archiválható, pedig az internet legfontosabb szegmense jelenleg. Ennek az előbb említett okok mellett a felhasználók által generált tartalom pusztán mennyisége a magyarázata. Magukon a szolgáltatókon kívül senkinek

² Dion Hoe-Lian Goh and Peng Kin Ng, „Link Decay in Leading Information Science Journals,” *Journal of the Association for Information Science and Technology* 58, 1. sz. (2007): 15–24, <https://onlinelibrary.wiley.com/doi/full/10.1002/asi.20513>; Frank McCown, Sheffan Chan, Michael L. Nelson and Johan Bollen, „The Availability and Persistence of Web References in D-Lib Magazine,” in *5th International Web Archiving Workshop and Digital Preservation* (2005), <http://arxiv.org/ftp/cs/papers/0511/0511077.pdf>.

³ *Wikipedia*, „Google Groups – Deja News,” hozzáférés: 2018.05.22, https://en.wikipedia.org/wiki/Google_Groups%2523Deja_News.

⁴ John Goerzen, „2007 Gopherspace Mirror,” film, hozzáférés: 2018.05.22, <http://archive.org/details/2007-gopher-mirror>.

⁵ Wayback Machine, Internet Archive, hozzáférés: 2018.05.22, <http://web.archive.org/>.

nincsen elég pénze és megfelelő technikája erre, ők viszont csak addig érdekeltek a megőrzésben, amíg az élő szolgáltatást nyereségesen tudják üzemeltetni. Így szűnt meg például a Microsoftnak a 2006-ban még 120 millió regisztrált taggal rendelkező Windows Live Spaces⁶ nevű blogplatformja 2011-ben, a Hyves⁷ nevű, 10 milliós létszámú holland közösségi oldal 2013-ban, valamint a magyar iWiW⁸ is 2014 június végén.

Hogy mekkora problémát jelent hosszú távú és kutatható archívummá alakítani a webkettes tartalmakat, arra jó példa az amerikai Library of Congress [Kongresszusi Könyvtár] esete a Twitterrel.⁹ 2010-ben a Library of Congress 2006 márciusáig, vagyis a Twitter indulásáig visszamenőleg megkapta az összes, mintegy 170 milliárd nyilvános *tweetet*, valamint az *élő folyamat* is, ami akkor napi 50 millió üzenet volt, ám ez a szám 2014-re már megtízszereződött, és azóta is folyamatosan nő. Mivel nemcsak a legfeljebb 140 karakterből álló szövegeket, hanem az azokhoz tartozó több mint százféle metaadatot is tárolni és indexelni kell, ezért a könyvtár – amúgy nem gyenge – számítógépes infrastruktúráján egy egész napig tartott volna egyetlen keresőkérdés lefuttatása. Magáncégek bevonásával 2014 közepére ígértek egy kísérleti szolgáltatást, de az azóta sem készült el.

1.1. Archívumfajták

A legtöbb internetarchívum jelenleg webarchívum, vagyis weboldalak vagy webhelyek valamilyen rendszeres vagy rendszertelen időközönként megismételt mentései. Egyre több a második generációs rendszer közöttük, amelyeket egy néhány éves üzemeltetés után alapjaiktól újraterveztek – és az első verzió honlapja jó esetben még megtalálható valamelyik webarchívumban. Céljuk és létrehozójuk alapján a főbb típusok a következők.

1.1.1. Magánarchívumok A cél lehet valamilyen érdeklődési körhöz vagy kutatási munkához való anyaggyűjtés, a hosszú távú hivatkozhatóság biztosítása, esetleg bizonyítékként való felhasználás. A magáncélú archiváláshoz felhasználóbarát célszoftverek (pl. HTTrack),¹⁰ böngészőkiegészítők (pl. Fireshot, ZipTabs), illetve ingyenes vagy fizetős online szolgáltatások és felhőtárhelyek (pl. Save Page Now, PageFreezer)¹¹ állnak rendelkezésre. De persze egy-egy weboldalt el lehet menteni magukkal a böngészőkkel is vagy a Zotero¹² nevű hivatkozáskezelő programmal, melyhez olyan modul (Hiberlink plugin for Zotero)¹³ is létezik, amely rögtön valamelyik nagy web-

⁶ Wikipedia, „Windows Live Spaces,” hozzáférés: 2018.05.22, https://en.wikipedia.org/wiki/Windows_Live_Spaces.

⁷ Wikipedia, „Hyves,” hozzáférés: 2018.05.22, <https://en.wikipedia.org/wiki/Hyves>.

⁸ Wikipedia, „iWiW,” hozzáférés: 2018.05.22, <https://hu.wikipedia.org/wiki/iWiW>.

⁹ Drótos László, „Michael Zimmer: A Kongresszusi Könyvtár Twitter archívuma,” recenzió (Michael Zimmer, *The Twitter Archive At the Library of Congress: Challenges for Information Practice and Information Policy* (2015)) *Tudományos és Műszaki Tájékoztatás* 62, 11–12. sz. (2015): 445–447, <https://tmt.omikk.bme.hu/tmt/article/download/610/581>.

¹⁰ HTTrack Website Copier, hozzáférés: 2018.05.22, <https://www.httrack.com/>.

¹¹ Wayback Machine, „Save Page Now,” hozzáférés: 2018.05.22, <http://web.archive.org/>; PageFreezer, hozzáférés: 2018.05.22, <https://www.pagefreezer.com/>.

¹² Zotero, hozzáférés: 2018.05.22, <https://www.zotero.org/>.

¹³ Hiberlink Zotero plugin, hozzáférés: 2018.05.22, <http://hiberlink.org/zotero.html>.

archívumba menti a megőrizni kívánt oldalt, és az onnan visszkapott archív URL-t is felveszi a Zotero adatbázisába.

1.1.2. Céges archívumok Az üzleti szférában egyre jellemzőbb az internetes tartalmak mentése. A motiváció a vállalat történetének megőrzése vagy egyszerűen az a törvényi előírás, hogy archiválniuk kell minden hivatalos kommunikációt az ügyfelekkel – beleértve a honlapjukon és a különböző webkettes csatornáikon közzétett tartalmaikat is. Előbbire jó példa a Coca-Cola,¹⁴ amelynek az archívuma hatmillió weboldalt őriz a cég különböző internetes felületeiről az 1995-ös első honlapig visszamenően. A saját anyagok mentése mellett a versenytársak vagy az adott üzletág online tartalmait is szokták gyűjteni statisztikai, adatbányászati, piac- és trendkutatási célokból. Több kulcsrakész, professzionális rendszer (pl. Q-Suite, Presurf)¹⁵ is kapható ma már, amelyekkel nemcsak webes dokumentumok, hanem videók, Skype-beszélgetések, mobiltelefonos üzenetváltások, üzleti tranzakciók (pl. a cég webshopjából) egyaránt rögzíthetők, időbélyeggel és digitális aláírással hitelesítve, hogy egy jogi vita esetén a bíróság is elfogadja őket bizonyítékként. Ilyen rendszereket SaaS (Software-as-a-Service) formában is lehet bérelni (pl. NetTrack, Cloud Preservation, Scrapinghub),¹⁶ ahol a megrendelőnek nem kell semmit telepítenie és tárolnia, hanem egy adminisztrációs felületen keresztül tudja ütemezni az archiválási feladatokat, és a lementett tartalom valamilyen felhőtárhelyen kerül megőrzésre.

1.1.3. Intézményi archívumok Közgyűjtemények (könyvtárak, levéltárak, múzeumok), egyetemek és kutatóintézetek, tudományos és civil szervezetek, kormányzati szervek egyaránt építenek alkalmi jelleggel vagy hosszabb távon webarchívumokat. Az egyik ok ezeknél is az intézmény történetének dokumentálása, vagyis a saját honlap és egyéb internetes felületek szisztematikus mentése. Az állami szervek esetében pedig több országban jogszabály írja elő, hogy elérhetőeknek kell maradniuk a korábbi, esetleg már érvényüket veszített, ezért az élő honlapról lekerült dokumentumoknak is. Emellett tematikus gyűjteményeket is építenek egyre több helyen: a városi könyvtárak például helyismereti, helytörténeti tartalmakat mentenek, a tudományos intézmények a kutatási területüknek megfelelő forrásokat, a civil szervezetek pedig az általuk képviselt ügy internetes lenyomatait. Az intézményi webarchívumok is készülhetnek a céges archívumoknál említett professzionális rendszerekkel, illetve fizetős felhőszolgáltatásokkal, de gyakoribb a nyílt forráskódú

¹⁴ Ted Ryan, „1s and 0s: The History of The Coca-Cola Company’s Website,” *Coca-Cola Company*, 2012. nov. 08., <http://www.coca-colacompany.com/stories/1s-and-0s-the-history-of-the-coca-cola-companys-website>.

¹⁵ Q-Suite, hozzáférés: 2018.05.22, <https://www.qumram.com/products>; Presurf, hozzáférés: 2018.05.22, <http://www.capsis.nl/en/websitearchiving/presurf/introduction/>.

¹⁶ NetTrack, hozzáférés: 2018.05.22, <http://www.capsis.nl/en/websitearchiving/nettrack/introduction/>; Cloud Preservation, hozzáférés: 2018.05.22, <http://www.nextpoint.com/>; Scrapinghub, hozzáférés: 2018.05.22, <https://scrapinghub.com/>.

szoftverekből összerakott saját megoldás (pl. NetarchiveSuite),¹⁷ valamint a nonprofit archiváló szolgáltatások (pl. Archive-It, archive.is, ArchivetheNet)¹⁸ használata.

Néhány érdekesebb projekt: York University Web Archives¹⁹ (az egyetem saját webhelyei, metaadatokkal együtt letölthető csomagokban is), Web Archive of Cacak²⁰ (a szerb Čacak város könyvtárának kis helyismereti gyűjteménye), CyberCemetery²¹ (az Egyesült Államok megszűnő kormányzati honlapjainak utolsó állapota), Human Rights Web Archive²² (a Columbia University Libraries által mentett emberi jogi témájú webhelyek), Contemporary Composers Web Archive²³ (54 modern zeneszerző – köztük Ligeti György – honlapjai az amerikai zenei könyvtárak szakembereinek válogatásában), Digital Archive for Chinese Studies²⁴ (az Universität Heidelberg Institut für Sinologie 2001 óta épített gyűjteménye), Latin American Web Archiving Project²⁵ (a University of Texas LANIC központjának gyűjtése latin-amerikai politikai pártokról és választásokról).

1.1.4. Nemzeti archívumok Rendszerint a nemzeti, állami könyvtár vagy az általa vezetett intézményi konzorcium tartja fenn ezeket, és a helyi kötelespéldányra vonatkozó törvény szabályozza a működésüket. A cél az adott nemzet digitálisan születő kultúrájának megőrzése a jövő számára. Ez történhet a nemzeti webtér időnkénti (éves vagy féléves) aratásával, illetve egy kellően reprezentatív, néhány ezer vagy néhány tízezer webhelyet tartalmazó részhalmaz gyakoribb (havi, heti vagy akár napi) mentésével. A legtöbb országban mindkettőt alkalmazzák, mert jól kiegészítik egymást: a teljes körű aratás egy átfogó pillanatképet rögzít, de ritkábban, a szelektív gyűjtéssel pedig gyakrabban és jobb minőségben lehet az érdekesebb, értékesebb tartalmakat archiválni, akár az országhatáron kívüli szerverekről is – és ezek kisebb számossága még azt is megengedi, hogy részletesebben metaadatolják, katalogizálják őket, ami a visszakeresést nagyban megkönnyíti. E mellett szokás még eseményalapú mentéseket is csinálni néhány napig vagy hétig, például valamilyen világraszóló rendezvény, választási kampány vagy természeti katasztrófa esetén, hogy a sajtóban és a közösségi fórumokon megjelenő információkból és reakciókból minél többet tudjanak rögzíteni. A szellemi tulajdont és személyiségi jogokat védő szabályok miatt a legtöbb országban csak helyben, a könyvtáron vagy a zárt könyvtári hálózaton belül

¹⁷ NetarchivSuite, hozzáférés: 2018.05.22, <https://sbforge.org/display/NAS/NetarchiveSuite>.

¹⁸ Archive-It, hozzáférés: 2018.05.22, <http://archive-it.org>; archive.is, hozzáférés: 2018.05.22, <http://archive.is>; ArchivetheNet, hozzáférés: 2018.05.22, <http://archivethe.net/en>.

¹⁹ York University Web Archives, hozzáférés: 2018.05.22, <https://digital.library.yorku.ca/yul-232039/web-archives>.

²⁰ Web Archive of Cacak, hozzáférés: 2018.05.22, <http://cacak-dis.rs/digital/english/web-archive-of-cacak/>.

²¹ CyberCemetery, hozzáférés: 2018.05.22, <https://govinfo.library.unt.edu>.

²² Human Rights Web Archive, hozzáférés: 2018.05.22, <https://hrwa.cul.columbia.edu>.

²³ Contemporary Composers Web Archive, hozzáférés: 2018.05.22, https://library.columbia.edu/bts/web_resources_collection/contemporary_composers_web_archive.html.

²⁴ Digital Archive for Chinese Studies, hozzáférés: 2018.05.22, http://www.zo.uni-heidelberg.de/boa/digital_resources/dachs/index_en.html.

²⁵ Latin American Web Archiving Project, University of Texas, hozzáférés: 2018.05.22, <http://lanic.utexas.edu/project/archives/>.

lehet hozzáférni a webarchívumhoz, dedikált gépekről, másolási lehetőség nélkül. De a szelektíven mentett webhelyekből – amelyeknél erre az eredeti tartalomgazda engedélyt adott – szoktak azért egy távolról is elérhető, böngészhető szolgáltatást is csinálni (ilyen pl. a szlovén webarchívum),²⁶ illetve a metaadatok általában a teljes gyűjtemény esetében nyilvánosak, és vagy egy külön adatbázisban (pl. a Library of Congress webarchívumának kereső- és böngészőfelülete),²⁷ vagy a könyvtár katalógusában visszakereshetők (pl. egy archivált webhely rekordja a svájci nemzeti könyvtár katalógusában).²⁸ A nemzeti könyvtárak többsége saját, nyílt forráskódú szoftverekből álló rendszert működtet a webarchiváláshoz (a katalán webarchívumhoz például ezeket a szoftvereket használják: Heritrix, Wayback, NutchWax, Wera, Web Curator Tool, Hadoop), gyakran egy IT-partner segítségével (pl. a szlovák nemzeti könyvtár a Tempest céggel),²⁹ de arra is van példa, hogy kiszervezték a feladatot, és egy professzionális webarchiváló céget bíztak meg vele (pl. az írországi web mentését és annak szolgáltatását a nemzeti könyvtár számára az *Internet Memory Foundation* végzi).³⁰ Az élő webről való aratás mellett létezik olyan megoldás is, hogy maga a tartalomszolgáltató küldi be valamilyen szabványos adatcsere-csatornán át a webhelyén megjelent új tartalmakat a könyvtári archívumba. A folyamatosan változó és robotokkal amúgy is nehezen bejárható, dinamikusan generált weboldalakból álló hírportálok esetében ez a legjobb megoldás.

Jelenleg körülbelül 40 projekt sorolható a nemzeti szintű archívum kategóriájába, de ez csak valamivel több mint 30 országot jelent, mert egyes tartományoknak, tagállamoknak vagy nagy nemzetiségeknek külön archívuma van. Néhány példa:

- PANDORA: ausztrál könyvtári konzorcium keretében működik 1996 óta, szelektíven mentenek, valamint katalogizálnak már közel 50 ezer webcímet, és az Internet Archive segítségével időnként a teljes .au domént is learatják.³¹
- LCWA: a Library of Congress 2000-ben – akkor még MINERVA néven – indított projektje, melynek keretében több mint 11 ezer webhelyet archiválnak, és eseményekről is csinálnak részgyűjteményeket, például a szeptember 11-i terrortámadás, a 2002-es téli olimpia, az iraki háború.³²
- UKWA: 2004-ben egy könyvtári együttműködés keretében létrejött brit webarchívum, amelynek három gyűjteménye van: egy több mint 15 ezer oldalból álló válogatott állomány, az Internet Archive-től átvett 1996–2013 közötti mentés

²⁶ Spletni arhiv, Narodne in univerzitetne knjižnice, hozzáférés: 2018.05.22, <http://arhiv.nuk.uni-lj.si>.

²⁷ Library of Congress, hozzáférés: 2018.05.22, <https://www.loc.gov/websites/>.

²⁸ Helveticat, Schweizerische Nationalbibliothek (NB), hozzáférés: 2018.05.22, <http://www.helvetica.ch/lib/item?id=chamo:1745898>.

²⁹ Central Archiving Platform, hozzáférés: 2018.05.22, <https://www.tempest.sk/products-and-services/central-archiving-platform-2d5.html>.

³⁰ Web Archive, National Library of Ireland, hozzáférés: 2018.05.22, https://www.nli.ie/en/web_archive.aspx.

³¹ PANDORA, Australia's Web Archive, hozzáférés: 2018.05.22, <http://pandora.nla.gov.au>.

³² Archived Web Sites, Library of Congress, hozzáférés: 2018.05.22, <https://www.loc.gov/websites/>.

az Egyesült Királyság webhelyeiről és az .uk címtartomány 2013 utáni saját mentései.³³

- WebArchiv: a cseh nemzeti könyvtár 2000-ben indult projektje, melyben selektív, eseményalapú és a teljes cseh webtérre kiterjedő archiválást végeznek. Eddig 5129 tartalomszolgáltatóval kötöttek szerződést.³⁴
- WARP: a japán National Diet Library 2002-től fejlesztett, már harmadik generációs webarchiváló rendszere, mellyel 2015-ben közel 11 ezer webhelyet mentettek, és ezekből 280 ezer fontosabb dokumentumot kigyűjtve önállóan is katalogizáltak.³⁵

1.1.5. Globális archívumok A korábban már említett, messze a legnagyobb Internet Archive mellett van még egy-két projekt, amelyek nem nemzet, földrajzi hely vagy téma alapján fókuszáltak. Ilyen például a Common Crawl nevű, kaliforniai székhelyű webarchiváló kezdeményezés,³⁶ mely 2011 óta ingyenesen letölthető és kutatható halmazokat gyűjt a nyilvános webről, jelenleg évi négyszeri aratással. 2015 végén már 1.82 milliárd weboldalt tettek így módon elérhetővé az Amazon felhő tárhelyéről. Ide sorolható még egy 2011–2013 közötti kísérleti EU-s projekt, a BlogForever.³⁷ Ennek keretében közel 210 ezer blogot mentettek és elemeztek ki azzal a céllal, hogy kidolgozzák ennek a műfajnak az archiválási technológiáját. Itt érdemes megemlíteni a 2009 óta létező, főként a veszélyeztetett internetes szolgáltatások megőrzésére szerveződött, Archive Team nevű – civilekből és szakemberekből álló – „akciócsoportot”,³⁸ illetve annak WikiTeam részét is,³⁹ amely eddig már 27 ezer önálló *wikit* és több *wikifarmot* mentett le az Internet Archive-ba.⁴⁰

A különböző nemzetközi, nemzeti és helyi internetarchiválási kezdeményezéseket egy 2003-ban a francia nemzeti könyvtár és 12 partnerintézmény által alapított konzorcium, az IIPC (International Internet Preservation Consortium)⁴¹ fogja össze, jelenleg 54 tagja van. A szervezet céljai: az internet megőrzésével foglalkozók közötti tapasztalatcsere, az ehhez szükséges technológiák közös fejlesztése, a szabványosítás. Külön munkacsoportok foglalkoznak a begyűjtés, a megőrzés, a hozzáférés, a ráépülő szolgáltatások és az oktatás témaival. Éves konferenciákat rendez, közös projekteket koordinál, szoftvereket fejleszt.

³³ UK Web Archive, hozzáférés: 2018.05.22, <http://webarchive.org.uk>.

³⁴ Webarchiv, the Museum of Czech Web, hozzáférés: 2018.05.22, <http://www.webarchiv.cz/en>.

³⁵ WARP: Web Archiving Project, hozzáférés: 2018.05.22, http://warp.da.ndl.go.jp/info/WARP_en.html.

³⁶ Common Crawl, hozzáférés: 2018.05.22, <http://commoncrawl.org/>.

³⁷ BlogForever, hozzáférés: 2018.05.22, <http://web.archive.org/web/20160729112149/http://blogforever.eu/>.

³⁸ Archive Team, hozzáférés: 2018.05.22, <https://www.archiveteam.org/>.

³⁹ WikiTeam, hozzáférés: 2018.05.22, <http://www.archiveteam.org/index.php?title=WikiTeam>.

⁴⁰ Internet Archive: WikiTeam, hozzáférés: 2018.05.22, <https://archive.org/details/wikiteam&tab=about>.

⁴¹ International Internet Preservation Consortium (IIPC), hozzáférés: 2018.05.22, <http://netpreserver.org/>.

1.2. A magyar helyzet

A hazai helyzet sajnos röviden összefoglalható: a kilencvenes évek második felében indított magyar webes keresők (Heuréka, Góliát, Altavizsla/Vizsla), majd a 2010 körül megjelent újgenerációs társaik (Bluu, Szörcs, Miner, PolyMeta/Johu, RichPOI) robotjai által gyűjtött magyar tartalomra nem épült webarchívum – ahogyan például a portugáloknál történt –, és azóta már el is tűntek ezek a rendszerek adatállományaikkal együtt, mert nem bírták a versenyt a Google keresőjével. Az MTA SZTAKI 2008–2013 között részt vett két európai uniós K+F-projektben (LIWA és LAWA),⁴² melyek a webarchiválás technológiájának megújítását és a webarchívumok kutatási célú felhasználásához szükséges módszerek és szoftverek kidolgozását célozták. A 2010-es évek elején az ELTE Tudománytörténet és Tudományfilozófia Tanszékének tudományometriai munkacsoportja végzett egy fókuszált webarchiválást.⁴³ Mintegy 400 magyar webhelyet: kutatóintézeti, valamint egyetemi és főiskolai honlapokat mentettek kéthetes periodicitással. A lementett tartalmat ki is elemezték olyan szempontból, hogy mit és mennyit kommunikálnak magukról online ezek az intézmények. A hazai könyvtári szférában 2006-ban hangzott el az első javaslat egy Magyar Internet Archivum (MIA) létrehozására.⁴⁴ A szándék az Országos Széchényi Könyvtár 2007-es munkatervébe is bekerült,⁴⁵ és bár több próbálkozás is volt a szükséges forrás megteremtésére (együttműködve a Szegedi Tudományegyetem könyvtárával, illetve az NIIF Programmal), ezek nem vezettek eredményre, így annak ellenére, hogy az OSZK-ban az egyedi internetes dokumentumok gyűjtése és feldolgozása már régóta folyik a MEK, EPA és DKA szolgáltatások⁴⁶ keretében, webhelyeket még nem archivál a nemzeti könyvtár. A 2017 elejétől 2018 végéig tartó OKR (Országos Könyvtári Rendszer) nevű projektbe viszont végre bekerült a webaratás tesztelése,⁴⁷ melyhez az infrastruktúrát a KIFÜ–NIIF⁴⁸ biztosítja. A tervek szerint néhány száz – főként kulturális és tudományos – webhely kerül többszöri lementésre, és lesz két kísérlet a .hu alá tartozó szerverek teljes körű aratására is. Az elsődleges cél egyelőre még csak a tanulás, a szükséges kutatási és fejlesztési munka elvégzése, egy üzemszerűen működő magyar webarchívum feltételeinek megteremtése.

⁴² MTA SZTAKI, Living Web Archives (LiWA), hozzáférés: 2018.05.22, <https://www.sztaki.hu/en/projects/liwa>; MTA SZTAKI, Longitudinal Analytics of Web Archive Data (LAWA), hozzáférés: 2018.05.22, <https://www.sztaki.hu/projektek/lawa>.

⁴³ Gulyás László, „Magyar Internet Archivum pilot és elemzés,” prezentáció (ELTE, 2014. április 14.), hozzáférés: 2018.07.18, <https://slideplayer.hu/slide/2647111/>.

⁴⁴ Drótos László, „Mi a MIA? Javaslat egy Magyar Internet Archivum létrehozására,” *Tudományos és Műszaki Tájékoztatás* 53, 6. sz. (2006): 267–274, http://tmt.omikk.bme.hu/show_news.html?id=4431&issue_id=473.

⁴⁵ Bibliotheca Nationalis Hungariae, „Az Országos Széchényi Könyvtár programja, 2007,” hozzáférés: 2018.05.22, http://www.oszk.hu/sites/default/files/szakmai_munkaterv_2007_0.pdf.

⁴⁶ Magyar Elektronikus Könyvtár, hozzáférés: 2018.05.22, <http://mek.oszk.hu/>; Elektronikus Periodika Archivum és Adatbázis, hozzáférés: 2018.05.22, <http://epa.oszk.hu/>; Digitális Képtár, hozzáférés: 2018.05.22, <http://dka.oszk.hu/>.

⁴⁷ Országos Széchényi Könyvtár, Magyar Internet Archivum, „OSZK webaratás – teszt fázis,” hozzáférés: 2018.05.22, <http://mekosztaly.oszk.hu/mia/>.

⁴⁸ Kormányzati Informatikai Fejlesztési Ügynökség, hozzáférés: 2018.05.22, <http://kifu.gov.hu/>; Nemzeti Információs Infrastruktúra Fejlesztési Program, hozzáférés: 2018.05.22, <https://niif.hu>.

Egy kultúrának a webtérben való szereplése és láthatósága ma már a globális verseny része, ezért a döntéshozóknak óriási a felelőssége, hogy az ezen a területen keletkező hiányosságaink, mulasztásaink ne okozzanak versenyhátrányt.

1.3. Technológia

A magán és a kisebb intézményi archívumoknál a már említett letöltőalkalmazásokat használják, melyekkel az ismerős Windows-környezetben vagy akár közvetlenül a böngészőből lehet weboldalt vagy webhelyeket lementeni. A cégek sokszor valamilyen kulcsrakész archiválórendszert vesznek meg, amely a webes tartalmak letöltése mellett API-kon (alkalmazásprogramozási felületeken át) tud menteni például levelezőrendszerekből és más kommunikációs csatornákról, webkettes platformokról vagy sugárzott médiafolyamokból, sőt akár a tranzakciós módszert is támogatja, vagyis amikor a webszerver minden olyan dokumentumból automatikusan elküld egy másolatot az archívumba, amelyet egy felhasználó lekért. A nagyméretű webarchívumok mind szoftveres robotokat, ún. keresőrobotokat (*crawler*) futtatnak. Ezek egy előre megadott URL-címlistából kiindulva derítik fel a weboldalak közötti linkeket, és az üzemeltető által definiált szabályok alapján döntenek el, hogy melyeket kövessenek, melyekről töltsék le az ott található weboldalt és a beléjük ágyazott egyéb fájlokat. A legtöbb nemzeti archívum már az Internet Archive által fejlesztett Heritrix crawlert⁴⁹ használja, amely nagyméretű, szabványos WARC-csomagokba⁵⁰ menti a megtalált digitális objektumokat. Ezekből később – az eredeti URL-címük alapján – a szintén ingyenes és nyílt forráskódú OpenWayback nevű szoftverrel⁵¹ rekonstruálhatók és nézhetők meg az archivált webhelyek különböző időpontbeli mentései. Természetesen le is lehet indexelni az archívumban levő szöveges fájlokat, és akkor a teljes szövegű visszakeresés is lehetséges. Erre a célra többféle szoftver is szóba jöhet, mint például a kifejezetten webarchívumokhoz kialakított NutchWAX.⁵² A letöltendő webhelyek nyilvántartásához, a mentések gyakoriságának, mélységének és egyéb paramétereinek beállításához, a begyűjtött anyag minőségének ellenőrzéséhez, a leíró metaadatok elkészítéséhez és az eredeti tartalomgazdától kapott engedélyek kezeléséhez szükség van még egy keretrendszerre. Egyes archívumok ezt maguk fejlesztik a kezdetektől fogva, de erre a célra is léteznek már nyílt forráskódú eszközök, például a Web Curator Tool.⁵³ Érdekes még megemlíteni a Memento Project⁵⁴ keretében kidolgozott megoldást, amely a webszerverek által használt HTTP-protokollt egészíti ki egy „Datetime” elemmel.⁵⁵ Ennek segítségével a kliens (pl. egy webböngészőt használó ember) egy

⁴⁹ Heritrix, hozzáférés: 2018.05.22, <http://crawler.archive.org/index.html>.

⁵⁰ Web ARChive (WARC) Format, hozzáférés: 2018.05.22, <https://iipc.github.io/warc-specifications/specifications/warc-format/warc-1.1/>.

⁵¹ Wayback, hozzáférés: 2018.05.22, <http://archive-access.sourceforge.net/projects/wayback/>.

⁵² NutchWAX, hozzáférés: 2018.05.22, <http://archive-access.sourceforge.net/projects/nutchwax/>.

⁵³ Web Curator Tool, hozzáférés: 2018.05.22, <http://dia-nz.github.io/webcurator/>.

⁵⁴ Time Travel Service, „Memento Guide – Introduction to Memento,” hozzáférés: 2018.05.22, <http://www.mementoweb.org/guide/quick-intro/>.

⁵⁵ H. Van de Sompel, M. Nelson, R. Sanderson, „HTTP Framework for Time-Based Access to Resource States: Memento,” *Request for Comments* (2013. dec.), <https://tools.ietf.org/html/rfc7089>.

weboldal aktuális állapota helyett le tudja kérni annak adott időpontbeli vagy ahhoz legközelebbi mentését a világ webarchívumainak valamelyikéből. Ezzel a megoldással a web olyan médium lett, amelynek múltja is van, nemcsak jelene.

1.4. Problémák

Annak ellenére, hogy az Internet Archive több mint 20 éve archiválja a globális webet, és számos országban már üzemszerűen működnek a webarchívumok, az internet megőrzésének feladata még messze nincs megoldva. Az egyik nagy probléma, hogy csak a weben van a hangsúly, és annak is leginkább a hagyományos változatait: honlapokat, blogokat, e-folyóiratokat, fórumokat stb. lehet jól aratni és visszaneézhetővé tenni a jelenlegi technológiával. Az internet igazán dinamikus szegmensei, a közösségi platformok, a stream formában sugárzott rengeteg médiatartalom, az üzenő- és csevegőcsatornák, a dokumentum- és fájlmegosztó szolgáltatások, a szerverek közvetítése nélküli peer-to-peer kapcsolatokon zajló adatcsere, a számtalan, saját API-n kommunikáló mobilapplikáció és persze az egész *deep* és *dark web* kimarad ezekből a webarchívumokból. Vannak ugyan ezeken a területeken is próbálkozások, ígéretesnek tűnő szoftverek, de komoly méretű archívumot és szolgáltatást még senkinek sem sikerült ilyenekből felépítenie – és a gyors technológiai változások és magas költségek miatt nem is nagyon van rá esély.

Persze a meglevő webarchívumoknak is örülni kell, mert így is hatalmas értéket képviselnek, nélkülük teljesen elveszett volna a 20. század végi és 21. század eleji történelem internetes leképeződése. De ahhoz, hogy ne pusztán digitális *raktárak*, hanem tudományos kutatásra is alkalmas *digitális könyvtárak*, *levéltárak* és *múzeumok* legyenek, még sok mindent meg kell oldani, ki kell fejleszteni. Ha kicsit kutakodunk például az Internet Archive Wayback Machine felületén, rövid idő alatt feltűnnek a rendszer korlátai. Például csak URL-cím alapján lehet keresni. (Igaz, az idén megjelent új verzióban már az egyes webhelyek kezdőlapjára mutató linkek szövege is kereshető, de az évek óta ígért teljes szövegű kereső még mindig várat magára, ami nem is csoda, mert a feladat tulajdonképpen nagyobb, mint a Google keresőjét működtetni, mert annak „csak” az éppen létező webről kell releváns találatokat adnia.) Sok mentésnél jön hibaüzenet; vagy a linkek nem követhetők, vagy a menük és a belső keresők nem működnek – ezek mind a jelenlegi archiválótechnológia korlátait jelzik. Az automatikus módszerek mellett emberi felügyelettel működő, témára és minőségre fókuszált – ezért jóval kisebb – webarchívumoknál természetesen jobb a helyzet, de még ezeknél is bele kell törődni, hogy alapvetően töredékes, nagyon szemetes, rosszul strukturált, kevésbé metaadatolt és hatalmas bithalmazokról van szó, ráadásul tele gyorsan avuló fájlformátumokkal, amelyek megjelenítése külön problémákat fog okozni a távolabbi jövőben. Tipikus *big data* tehát egy webarchívum, amelynek a kutathatóvá, bányászhatóvá és vizualizálhatóvá tételéhez a már meglevő eszközök mellett még sok mindent ki kell fejleszteni. Hogy lesznek-e új generációs, a mainál jobb technikák az internet megőrzésére, a megőrzött tartalom elemzésére és feldolgozására, az jelentős részben attól függ, hogy a humán- és társadalomtudományok szakemberei mennyire igénylik ezeket, milyen innovatív kutatási módszereket találnak ki az internetes archívumok anyagának hasznosítására.

2. Az internet és a történelem

Az internet viszonylag korán került a történeti kutatók látókörébe, természetesen először kommunikációs közegként, azután már a különféle források és irodalmak tárolásának és elosztásának módjaként is. S nem szabad megfeledkezni arról sem, hogy a nagyon korai időszakban fellelhető online könyvtárkatalógusok mennyire fontos tájékoztató pontot jelentettek. Érdekes, hogy itthon már a '90-es évek végén több kiadvány tárgyalta, illetve leírta a történész és a hálózat viszonyát,⁵⁶ a rengeteg külföldi hasonlóról nem is beszélve.⁵⁷ Az interneten fellelhető történeti források listáját és elérhetőségét ma már nem is lehet nyomtatott kiadványokban összegyűjteni.

A digitális bölcsészeti kutatások egyik legfontosabb területe a *történeti* lett, most erősen hangsúlyozva e szóhasználatnak a *történettudományin túlmutató* hatókörét. Az pedig magától értetődő, hogy a digitális bölcsészet számára a webarchívum a *jövő levéltára*, amelynek sokrétű kutatása nyomán a megismerési folyamatban új és új szempontok, módszerek, információs rétegek és persze eredmények tárulnak majd fel.⁵⁸

2.1. A megőrzés biztonsága

Mindnyájunk tapasztalata a webes információ ilyen vagy olyan okokból való eltűnése. A webtér ingatagsága még a régi könyvek fennmaradáshoz képest is riasztó, a Vizsolyi Biblia mai fellelhetőségének valószínűsége 400 év távlatából is nagyobb, mint egy átlagos weboldalnak.⁵⁹ Ismertek példák arra is, amikor véletlenül vagy nemtörődomségből, de olykor nyilvánvaló tudatosság okán a mai politikai és gazdasági történések dokumentumai, jelenkori tudásunk forrásai tűnnek el a hálózatról.

A *The Web as History* című kötetben több példát hoznak a szerkesztők a *tudatos eltüntetésre*. 2013 végén fordult elő az a sajnálatos eset, hogy a brit Konzervatív Párt szervereiről törölték azokat a korábbi tartalmakat, amelyek a párt és David Cameron

⁵⁶ Sennyei Pongrácz, *A hálózat használata a történettudomány területén*, NIIF Információs Füzetek 1 (Budapest: NIIF, 1998). E sorozat több mint 20 füzetéből (szerk. Drótos László és Kokas Károly) több másnak is vannak történeti referenciái, mint pl. a filozófiai, irodalmi és nyelvi vagy ókortudományi kiadványnak. A sorozat online fellelhetősége: <http://mek.oszk.hu/01200/01280/html/>; lásd még Komáromy Gábor, *Történelem az Interneten* (Budapest: Kossuth Kiadó, 1998).

⁵⁷ Ez utóbbiakról ad viszonylag korai képet a Daniel J. Cohen és Roy Rosenzweig által szerkesztett *Digital History: a Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, 2006). Az egyik legkorábbi és máig hivatkozott összefoglaló: Andrew McMichael, Roy Rosenzweig and Michael O'Malley, „Historians and the Web: A Beginner's Guide,” *Perspectives on History* (1996 jan.), <https://www.historians.org/publications-and-directories/perspectives-on-history/january-1996/historians-and-the-web-a-beginners-guide>, <https://www.historians.org/publications-and-directories/perspectives-on-history/january-1996/historians-and-the-web-a-beginners-guide>.

⁵⁸ A digitális bölcsészet és a történeti kutatások általános helyzetképéről lásd Kokas Károly, „Digitális bölcsészet 2016: A bölcsészek és az informatikai megközelítés régen és most,” in *MONOKgraphia: tanulmányok Monok István 60. születésnapjára*, szerk. Nyerges Judit, Verók Attila, Zvara Edina (Budapest: Kossuth Kiadó, 2016), 405–412, <http://publicatio.bibl.u-szeged.hu/10296/>.

⁵⁹ Horváth Iván is figyelmeztetett erre, lásd „A hálózat hátránya: fennmaradás helyett pusztulás?” in Horváth Iván, *Magyarok Babelben* (Szeged–Budapest: JATEPress–Gépeskönyv, 2000), <http://magyar-irodalom.elte.hu/babel/2450.htm>.

számára kellemetlenné váltak. A politológus kutatók végül a törölt dokumentumokat a British Library webarchívumában találták meg. Az orosz-ukrán konfliktusban 2014-ben lelőtt utasszállító gép kapcsán „gondos kezek” szerették volna eltüntetni azt az orosz katonai körökből származó internetes hírt, miszerint orosz szakadárok lőtték volna le a gépet. De a bejegyzést az Internet Archive megőrizte a kutatóknak.⁶⁰

Ezek a példák azt illusztrálják, hogy a webtér információi sorsdöntő ügyekben hiányozhatnak vagy előkerülhetnek, s ebben a kérdésben a webtér archiválásának döntő jelentősége van. Az esetek arra is figyelmeztetnek, hogy ebben a helyzetben sincs másképp, mint a hagyományos történeti kutatás anyagainál: bár úgy látszik, aminek nincs lenyomata, az nem is létezett, az előfordulhat, hogy megtalálható jobb keresési módokkal vagy más archívumokban.

Más szempontból különösen fontos e területen, hogy a *nyílt adat* (open data) mentalitás és gyakorlat uralkodóvá váljon. Ez az alapelv meghatározó módon elvárja az adatok *elérhetőségének* és *hozzáférhetőségének* a megvalósítását, az *újrafelhasználás* és *továbbterjeszthetőség* feltételeinek rendezettségét és a feldolgozásban az univerzális részvétel lehetőségét, amely *a lehető legkevesebb korlát* felállítását engedi csak meg. Mindennek egyfajta következménye a nagyon magas fokú *interoperabilitás*, amely a különböző adathalmazok együttműködésének vagy vegyíthetőségének lehetőségét is jelenti, hisz ez teszi lehetővé a *különböző komponensek együttműködését*. Mindez nem csupán az adatbiztonsággal, az elérés és kutathatóság demokratizálásával függ össze. Ez a gondolkodás maximálisan támogatja a kutatások, az adatbányászati módszerek legteljesebb használatának lehetőségét is.⁶¹

2.2. A webtér mint történeti forrás⁶²

Az a gondolat, hogy maga a teljes internet forrása a történeti kutatásoknak, valamivel később, illetve párhuzamosan kerül elő a webarchiválási projektek értelmének és szükségességének indoklásaiban.⁶³ Úgy tűnik, a webtér és a történész viszonyának legalább négy vonatkozása van:

- a kifejezetten történeti kutatás számára készült repertóriumok, adatbázisok és más szolgáltatások;
- ezek a webarchiválás részeként mentésre kerülnek, így másodlagos elérhetőséget is ad az adott webarchívum, illetve gondoskodik arról, hogy az eredeti anyag ne tűnjön el a webtérből;

⁶⁰ Niels Brügger and Ralph Schroeder, „Introduction: the Web as History,” in *The Web as History: Using Web Archives to Understand the Past and the Present*, eds. Niels Brügger and Ralph Schroeder, (London: UCL Press, 2017), 1–2.

⁶¹ Vö. *Open Data Handbook*, hozzáférés: 2018.05.22, <http://opendatahandbook.org/guide/hu/what-is-open-data/>.

⁶² A kérdés legújabb és legátfogóbb vizsgálata a fent idézett *The Web as History* kötetben. Ez a szerkesztett kötet az első monografikus igényű kiadvány, amely arra összpontosít, hogy miként lehet a múlt archivált webtartalmait felhasználni a társadalom fejlődésének széleskörű kutatásához.

⁶³ Minderről általában és összefoglalóan: Niels Brügger, „Web History and the Web as a Historical Source,” *Zeithistorische Forschungen/Studies in Contemporary History, Online-Ausgabe* 9, 2. sz. (2012): 316–325, <http://www.zeithistorische-forschungen.de/2-2012/id=4426>.

- mindennek egy külön vonatkozása, hogy a periodikusan lementéseket tartalmazó webarchívum rétegeiben az *online történelmi kutatások eszköztárainak historiográfiája* is kutathatóvá válik;
- a *webtér archiválása*, amikor az internet maga válik történeti forrássá.⁶⁴

Dolgozatunk szempontjából természetesen most ez az utóbbi mód és út a lényeges. A web negyedszázada kezdte el átfogni, befogadni és befolyásolni az életünket. Ez a folyamat olyan gyorsan zajlott le, és a trend növekedése olyan intenzív volt, hogy már jelen állapotában szinte teljes körűnek mondható. Nyilvánvaló, hogy aki az elmúlt 25 év történetével – bármilyen aspektusú történetével – kíván foglalkozni, az nem kerülheti el a webtér vizsgálatát.

2.2.1. A szakirodalom, a módszertan és a forrástípusok A webarchívumok történeti kutatásokban való felhasználásának a tapasztalata értelemszerűen nem túl nagy. Bár számos példa van már arra, hogy történeti, irodalomtörténeti stb. tanulmányok hivatkozásában találunk webarchívumos adatot (lásd az alábbi fejezetekben), az archivált webtér mint forrástípus módszertani megközelítésének irodalma és a téma elméleti feldolgozása még kezdeti állapotban van. Ha rákeresünk a problémát legjobban megragadó *web historiography* kulcsszóra, ma még alig találunk többet, mint pár tucat cikket és néhány könyvet. Kis túlzással a hivatkozott szakirodalom felét-harmadát a téma legaktívabb kutatója, Niels Brügger (aki talán a *webhistoriográfia* szakkifejezés első alkalmazója is lehet) írta, szerkesztette vagy inicializálta, aki a dán Centre for Internet Studies vezetője és az Aarhus Egyetem professzora.

A szakirodalmi feldolgozás hiányosságainak – azon túl, hogy a webaratások és projektek kiteljesedése alig több mint egy évtizedes múltra tekint vissza – az is oka, hogy ennek a forrástípusnak az elemzése és jellegzetességeinek a feltárása valószínűleg a legkomplexebb forrásismereti kihívás a kutatóknak.

Ennek okai – többek között – a következők lehetnek:

- a vizsgálatok óriási terjedelmű adatokra vonatkozhatnak (*big data*);
- rengeteg más, korábban már létező, de egymással föltétlen össze nem kapcsolt egyéb forrást kívánunk egyben kezelni;
- minden létező, elektronikusan egyáltalán reprodukálható és megragadható médiatípus része lehet a fájlsomagnak;
- az adattömeg belső rendje és metaadatolása nagyon különféle lehet, s annak metódusához még nincs kialakult gyakorlat;
- a megjelenő és a kívánatos módszerek a történészi körökben még kevésbé ismert és naponta újdonságokat felmutató mesterséges intelligencia (MI, angol rövidítés AI) határmezsgyéjén mozognak;

⁶⁴ Érdeemes ehhez referálni a Web Archives for Historical Research (WAHR) csoport weblapját, ahol így fogalmaznak: „This project is among the first attempts to harness data in ways that will enable present and future historians to usefully access, interpret, and curate the masses of born-digital primary sources that document our recent past.” WAHR, hozzáférés: 2018.05.22, <https://uwaterloo.ca/web-archive-group/about>.

- ha a *webhistoriográfia* a történettudomány segédtudományává válik majd, különös jellemzője lesz, hogy több tudományterületet maga is segédtudományként használ (az informatika, könyvtártudomány, matematika, szociológia, statisztika és szociálpszichológia biztosan előkerül ebben a kontextusban);
- különlegessége a forrásnak, hogy látványosan a jövőnek készül, hiszen úgy gondolják a készítőik, hogy a pusztá elmentése a webtérnek a benne foglalt primer információkon túl később a fejlettebb módszerekkel egyre inkább és sokkal hatékonyabban kutatható lesz.

A fentieket és a hozzájuk hasonló szempontokat, megfigyeléseket most érdemes felvetni. Ezek felderítéséhez és igazi megoldásához azonban az apró részleteken keresztül vezet az út, vagyis részletes esettanulmányok sokaságának elkészítésével.⁶⁵

2.2.2. A teljesség és a rész A webarchiválás kezdeteitől világos tehát, hogy mindennek a teljes kultúra történetére és annak kutatására vonatkozó aspektusa is van. Különösen felerősödött ez a gondolat akkor, amikor később nyilvánvalóvá vált, hogy a politikai történések jelentős része a weben zajlik, és az összes többi médium is azt referálja. A szubkultúrák szinte teljesen az internetre költöztek, nem beszélve arról, hogy a fiatalok majdnem kizárólagos globális kommunikációs környezete a 21. század elejétől maga a webtér lett (YouTube, Facebook, Twitter stb.), hogy aztán az évszázad második évtizedével már a fejlett világ általános kommunikációs közművévé váljon.

A jövőben az archivált web várhatóan fontos szerepet kap majd a média- és a kommunikációtörténetben, de azon történészek, politológusok stb. számára is nélkülözhetetlen lesz, akik a közelmúlt folyamatainak mélyrétegeit vizsgálják. Mindebből következik továbbá, hogy érdemes vizsgálni az archivált webtér történeti forrásként való kezelésének sajátosságait, valamint azt, hogy milyen kihívások elé állítja a történetst az új médium, amelyet forrásként kíván kutatni.⁶⁶

A mentett webtömeg (akár óriás projekteket, akár tematikus vagy nemzeti programokat nézünk) méreténél és átfogó természeténél fogva speciális terepet ad a kutatásnak. Maga a pusztá adatmennyiség és az a tény, hogy az egészében is indexelt (ellentétben például azzal, hogy a levéltárak és azok együttese nem kutatható egészében digitálisan), különös lehetőségeket ad, de meg is rettentí a felhasználót. A webtérnek is megvannak a valósághoz képest a maga torzításai, amelyek hamis illúziót kelthetnek, hiszen az, hogy mi, mikor és milyen mértékben kerül a webre, nem föltétlen függ össze a dolog súlyával és fontosságával. Például a globális webtérben föltétlen vizsgálendő és figyelembe veendő az, hogy itt mindig is egyértelmű volt az angolszász kultúra, s különösen az angol nyelv dominanciája. Ezen belül a folyamat elején – az internet amerikai eredete miatt – az amerikai kulturális hatás egészében befolyásolta a webtermést

⁶⁵ Vö. Josh Cowls, „Cultures of the UK Web,” in *The Web as History: Using Web Archives to Understand the Past and the Present*, eds. Niels Brügger and Ralph Schroeder (London: UCL Press, 2017), 220–237. Itt a .uk tartomány alá eső mintegy 65 terabájtos mentés módszertani elemzéséről van szó, hogy az a lehető legjobban kutathatóvá váljon. A British Library, a University of London Történeti Kutatóintézetének és a University of Oxford Internet Intézete által vezetett Big UK Domain Data for Arts and Humanities (BUDDAH) projekt eredményeit mutatja be ez a dolgozat.

⁶⁶ Brügger, „Web History,” 316–325.

(például a franciák és a francia nyelvű anyag később érkezett és más prioritásokkal), míg a jelenben éppen a kínai webkultúra erőteljes térfoglalása zajlik, ami az egész webtér arányait is megváltoztatja. De ez nem csak a nyelvi kultúrákra igaz: vannak jellemzően hamar webre kerülő témák, intézmények, és vannak olyanok, amelyek ezen a téren sokkal visszafogottabbak (hasonlítsuk össze például a filmkultúra és az egyházi tartalmak webfoglalásának tér- és időszerkezetét). Ezeknek a szempontoknak a jelenkor-történeti webkutatásoknál való felvetése és bevitele külön lehetőségeket és persze veszélyeket rejt.

Ismert mondás, hogy a jövő generációi számára az a létező, amit a webtérben fellelhetünk. De a történésznek máshogy illik gondolkodnia, mert nem a webtartalmak történetét kutatja valójában, hanem a valóság lenyomatait keresi. Így számolnia kell azzal is, ami nincs, vagy ami abból a korszakból még nem hagyott lenyomatot a hálózaton. El kell különíteni az illúziókat a valóságos arányoktól.

E belső aránytalanságokat növeli és növelheti, hogy a már elkészült és működő webarchívumok kutathatósága sem egyforma: különféle személyiségi és szerzői jogi megfontolások korlátozhatják. A hozzáférés pusztá kényelmének is van kutatást torzító hatása.⁶⁷

2.3. Forrástípusok és esetek

Nem képzelhető el a lementett webtér jó történeti hasznosítása anélkül, hogy magának az anyagnak a természetéről ne alkotnánk pontos képet. Ez is különleges tudást követel, nyugodtan betehetjük hát a webtérarchiválást a kronológia, a pecséttan, az oklevéltan mint történeti segédtudományok közé. A mai kutatások ezen a téren jószerivel csak példákat produkálnak, néha sporadikusan, de a kísérletezés veleje, hogy megtaláljuk és kialakítsuk a megfelelő módszertant, meglegyenek azok a szempontok, amelyek a kutatás e speciális szegmensében érvényesíthetők.

Természetesen a történeti forráskezelésben feltett kérdések sorra következhetnek: mi pontosan a forrás? Milyen típusú „objektum” az? Annak egészéről vagy részéről beszélünk? Ki hozta azt létre? Köthető-e pontos időponthoz? Mi annak a hitelessége? stb. Látható, hogy a szokásos gondok a webtér történeti kutatásában sokszor egészen másképp vetődnek fel, olykor még az is előfordulhat, hogy a századokig természetesen tartott forráskezelési szempont nem érvényes vagy egyenesen értelmezhetetlen. Az a problémák egészen különleges vonatkozása e kontextusban, hogy számtalan hivatkozást kapunk és kaphatunk számunkra elérhetetlen vagy éppen elveszett forrásokra, gondoljunk csak az anyagban beágyazott linkekre, amelyek *nem* részei egyik archívumnak sem, illetve a rengeteg, akár elvileg a webtérben lévő, de magánjellegű szövegre (elsősorban a levelezésekre).

A másik szokatlannak tűnő szempont, hogy az analóg világ forrásaival ellentétben itt sokszor *elvileg* rendelkezhetünk az adott webdokumentum mások által való felhasználásának adataival is, tehát azzal, hogy számszerűsíthetően az adott forrást hányan és mikor használták. Mekkora hatást gyakorolhatott? A gondolat, történés recepciótörténete így a jövőben egészen új aspektust kaphat, mert a részben vagy

⁶⁷ Vö. Niels Brügger and Ralph Schroeder, „Introduction: the Web as History,” 1–17.

egészen alátámasztott vélelmek, feltételezések helyett nagyon konkrét hivatkozásokkal számolhat. Mindennek persze elsődleges feltétele az, hogy a webarchiválás és annak metainformációs rendszere és struktúrája alkalmas legyen ilyen vizsgálatokra.

Magáról a webes információ természetéről is érdemes megfontolásokat tenni. Minden történész tudja, mi a különbség – egy adott témában – a levéltárban feltárt jegyzőkönyv, egy egykori magánlevél, egy újságban megjelent vezércikk vagy riport, illetve egy politikusi nyilatkozat között. Ismereteink vannak ezekről a forrástípusokról, tudjuk őket azonosítani, s értékük és jellegük szerint kezeljük őket. Vajon megvan-e ez a biztonsága a történésznek a hálózati források esetében? Akkor, amikor azok nem az analóg kultúra pusztá digitális lenyomatai, hanem egészen új típusú források, olyanok, amelyek már a hálózat világában születtek, mint például a blogbejegyzések vagy egy cikk kommentjei.

S természetesen külön módszertani probléma az, hogy a webarchiválás konténerszerű, magában foglaló természete révén a legkülönbébb digitális objektumok szerepelhetnek az éppen használt archív anyagban, amelyeknek mind megvannak a tartalmi és formai sajátosságaik, sokszor nemcsak értelmezési, de műszaki probléma elé is állítva a kutatókat. (Az magában külön téma lehetne, hogy ezen médiumtípusok némelyike igen kicsiny múltra tekint vissza, ezért az elemzés is gyerekcipőben jár. Gondoljunk például a mobileszközök révén elszaporodott videódokumentumokra vagy a digitális tudás szélesülésével elkövetett – vidám vagy komoly – hálózati hoaxokra, továbbá az ún. memkultúrára, amely magában is a politika s később a történelem nagyon különleges *karikatúramozzanata* lesz.)

2.3.1. A keresés mint a kutatás maga A szakirodalomban már kirajzolódik, hogy ezen új típusú és iszonytató tömegű forrás kapcsán átértelmeződik a *keresés* fogalma is, hiszen a jól keresni tudás a magas szintű kutatásnak egyik alapfeltétele lesz.⁶⁸ A holland kutatók (egy egyetem, egy kutatóintézet és a nemzeti könyvtár együttműködésében) a *search as research* gondolathoz metodológiát és a különféle lehetséges forgatókönyvekhez példákat is publikáltak. Külön kiemelik, hogy vizsgálatuk célja elsősorban a keresőmotorok algoritmusainak tanulmányozása, melyek a felhasználók számára valójában rejtett módon rendezik és rangsorolják a találatokat (például a látogatottság, vagy az IP-címtartomány földrajzi közelsége szerint). Mindezt az ún. lekérdezés-tervezés, illetve a keresési eredmények összehasonlító elemzése azt a célt szolgálja, hogy a különféle keresőmotorok algoritmusát (a webtörténész) a nyilvánossághoz közelebb hozza, és azokat átláthatóvá tegye. A kereshető webarchívumok esetében maga az egész archívum is egy ún. analitikai egység lehet, amelyet globális *big data* metódusokkal is lehet értelmezni, ahol a szövegbányászati algoritmusok (mint például az egyre gyakrabban emlegetett *n-gram* feldolgozás), a mesterséges intelligencia bevonása (statisztikai alapú elemzések), és persze a grafikonmegjelenítés és -elemzés is fontos eszköz lehet.

Nem tárgya most dolgozatunknak, de felvethető, hogy a webarchívumon kívül a valóságos webkeresésben ugyanezek a szempontok, a keresőmotorok által előállított

⁶⁸ Anat Ben-David and Hugo Huurdeman, „Web Archive Search as Research: Methodological and Theoretical Implications,” *Alexandria* 25, 1–2. sz., 93–111.

eredmények milyen viszonyban vannak a teljes valósággal? És mindez vajon milyen mennyiségben és minőségben befolyásol mindent, amit az internetes keresési eredmények következtében teszünk és visszatermelünk a webtérbe, hogy aztán az, legalábbis részben, kvázi manipulált webtörténelemmé váljon? (E ponttól – a szövegvilágok befogadásában is – eljuthatunk akár Heisenberg nevezetes határozatlansági relációjának *bölcsész applikációjához*, a Luhmann alkotta tézishez, ti. a megfigyelő és megfigyelés torzító hatásmechanizmusának leírásához, hiszen ebben az esetben és gondolatmenetben lényegében ugyanarról van szó.)⁶⁹

2.3.2. Esettanulmányok és példák A történeti webkutatás fejlődésének jelen szakaszában a fenti és a fentihez hasonló elméleti és módszertani megfontolások mellett nagyon nagy szükség van jól kidolgozott és alaposan dokumentált esettanulmányokra. Messze vagyunk még attól, hogy mindez rutinná váljon, és ameddig a témán belüli specializáció le nem zajlik, addig fontos, hogy minden érintettet (majdnem) minden kutatás érdekeljen. Ennek oka, hogy ezen kutatásoknak a megismerése intuitív lehet: ötleteket és módszertani tanácsot adhat, új területekre, friss gondolatokra hívhatja fel a figyelmet, s az utánzásnak és az analógiáknak nagy jelentősége lehet. Ilyen szempontból is óriási haszna van az említett összefoglaló kötet esettanulmányainak, illetve a weben egyre szaporodó, tárgyunkba vágó iniciatíváknak.

Az ismert és tekintélyes webarchívumok anyagának hasznosításával tehát sorra születnek az elemzések, és ezek – madártávlatból nézve – legfőbb szembeűnő jellegzetessége a fantasztikus változatosság. Már most meglepő, hogy hányféle módszer és megközelítés lehetséges, s mennyi különleges metszete rajzolható ki az anyagnak. Vannak már nagyobb és átfogóbb, hagyományos történészi megközelítésben gondolkodó feldolgozások is, de talán érdekesebbek azok, amelyek különleges és friss szempontokat vetnek fel.

Egy egészen friss tanulmány a webarchívumok fontosságáról a humán tudományokban azt boncolgatja, hogy szinte minden archivált információ fontos lehet történelmi szempontból. Sőt, a kutatási szempontok mindig újabb és újabb vetületét mutathatják az anyagnak, mindez csak attól függ, hogy milyen kontextusban folyik éppen a kutatás. Például egy ingyenes hirdetés, amelyben egy használt gyerekkerékpár képe szerepel, első látásra bagatellnek tűnhet, és legföljebb az általános megőrzés szempontjából fontos. Azonban húsz év múlva kiderülhet, hogy ez a híres kerékpár-bajnok első kerékpárja, vagy érdekes lehet – hasonlókkal együtt – egy olyan kutató számára, aki éppen a kerékpártervezés technológiai fejlődését kutatja. Ez a példa jól mutatja, hogy milyen nehéz egy webarchívátornak eldönteni, mi a fontos és mi nem a jövő szempontjából.⁷⁰

Az idősorokban és a földrajzi vonatkozásokban (*georeferenciák*) is nagy potenciál van. Az említett tanulmány hangsúlyozza ezt, és példaként bemutatja, hogy a 2003-ban archivált portugál választási eredmények állami kezelésű honlapjai hogyan alkotnak az archívumban változatosan vizsgálható idősort egészen 1997 óta. Mindennek

⁶⁹ Vö. Szajbély Mihály, *A nemzeti narratíva szerepe a magyar irodalmi kánon alakulásában Világos után* (Budapest: Universitas Kiadó, 2005). A könyv az első magyar nyelvű kísérlet a Niklas Luhmann által kidolgozott rendszerelmélet adaptációjára a kifejezett bölcsészettudományok területén.

⁷⁰ Daniel Gomes and Miguel Costa, „The Importance of Web Archives for Humanities,” *Journal of Humanities & Arts Computing* 8, 1. sz. (2014), 113–114.

következtében a 2002-es választások eredményeinek vizsgálata és webprezentációs története is más megvilágításba került.⁷¹

De nem csupán a történelmi tények tekintetében fontos a webtér kutatása. Egy ausztrál vizsgálatban a hosszan elnyúló abortuszvitát illetően például érdekesebbnek tűnik maguknál az elsődleges eredményeknél az a szempont, hogy a *hálózaton való megtalálhatósága* a különféle álláspontokat közvetítő weblapoknak (az ún. *pro-choice* és *pro-life* érvelésnek) mennyire befolyásolhatta az álláspontok és vélemények alakulását.⁷² Így azután a keresési metódusok, a különféle keresőszolgáltatások fontosabb szereplőként tűntek fel, mint azt korábban gondolták. (Kicsit hasonló ez ahhoz, hogy az adott hír milyen nézettségű televíziós csatornán vagy hogy melyik időszámban látható.)

A tartalomelemzés, a szövegbányászat és annak teljes módszertana egyrészt jóval régebbre nyúlik vissza, mint a webarchiválás, másrészt még maga is fejlődésének kezdeteinél tart, természetesen rengeteg eredménnyel és nagyon sok projektben. Nem nehéz elképzelni, hogy ezek a metódusok, a konkrét szoftverek és a hozzájuk társuló gráfos és grafikonos kimenetelű vizualizációs eljárások, a mesterséges intelligenciát felmutató programok rendre megtalálják illetve egyre jobban felfedezik a webarchívumok lehetőségeit is. Ez a folyamat még az elején tart, de kétségtelenül az egyik legnagyobb kutatási potenciállal és perspektívával rendelkező terület.⁷³

3. Összefoglalás és a jövő

A nemzeti és intézményi webarchiválás minden egyes eleme fontos, de mint a Library of Congress 2001-es (!) webmegőrzésre irányuló jelentése is kimondta, ajánlatos a jövőnek való megőrzés biztonsága érdekében, hogy több nemzeti és más intézmény végezze a webtér lementését. A jövőben kiemelt fontosságú lesz, hogy ezek a projektek – akár világméretűen is – koordinálva legyenek, tervezésüket, működésüket a lehető legjobban összehangolják.⁷⁴

Elmondhatjuk, hogy a webarchiválás a kezdeteinél tart, tudományos hasznosítása pedig sokszor még tapogatózás jellegű. Az azonban világos, hogy a jelenkortörténet kutatásának nemsokára fő eszköze lesz. Olyan lesz számukra a webhistoriográfia, mint például a középkorászoknak az oklevéltan. S valljuk be, aki az archontológiában,

⁷¹ D. Gomes and M. Costa, „The Importance of Web Archives for Humanities,” 111.

⁷² Robert Ackland and Ann Evans, „Using the Web to Examine the Evolution of the Abortion Debate in Australia, 2005–2015,” in *The Web as History: Using Web Archives to Understand the Past and the Present*, eds. Niels Brügger and Ralph Schroeder (London: UCL Press, 2017), 186–187.

⁷³ A sok közül minderre jó példa lehet az alábbi, ahol nem webarchívumokban, hanem egy tudományos folyóirat archívumában próbáltak szoftveres elemzést prezentálni, de nyilvánvaló, hogy hasonlókat webaratásokban is el lehet képzelni. J. Comins and L. Leydesdorff, „RPYS i/o: Software Demonstration of a Web-based Tool for the Historiography and Visualization of Citation Classics, Sleeping Beauties and Research Rronts,” *Scientometrics* 3 (2016), <https://arxiv.org/ftp/arxiv/papers/1602/1602.01950.pdf>.

⁷⁴ „Many organizations and individuals will collect and preserve materials independently of any relationship to the Library of Congress. Indeed this independence is important for preservation. If different organizations, in different countries with different cultures, carry out different preservation programs the materials that they collect are vulnerable in different ways. This diminishes the risk of everything being lost through a single disaster or mistakes of technology and organization.” William Y. Arms, „A Report to the Library of Congress: Web Preservation Project Final Report,” Library of Congress, 2001. jún., <https://www.loc.gov/webarchiving/include/webpresf2.pdf>.

diplomatikában, heraldikában és társaiban nem jártas valamennyire, nem is lehet igazi középkorkutató. A webarchívum történeti hasznosítása azonban magára a történeti kutatásra is visszahat, hiszen olyasmiket tudhatunk általa – rejtett trendeket vagy egyáltalán megőrzött források sokaságát, a nagy időszeletek kutatásának lehetőségét stb. –, amelyek megsokszorozzák a történeti anyag mennyiségét, reflektálhatóságát és egészen új szempontokat vetnek fel.

A szerzőknek, akik könyvtárosok, engedjék meg még egy szempont: a struktúra és a tételek feltárása a webarchívumban. Köztudott, hogy a hagyományos levéltárak, kéziratárak és könyvtárak tudományos (pl. történészi) használatának színvonala mennyire függ az anyag rendezettségétől és feltártságától. Ez ránk, könyvtárosokra (levéltárosokra, archivátorokra) nagyon nagy felelősséget ró, hiszen a metaadatolás kulcskérdés lesz.⁷⁵ E folyamatokban jól látható, hogy az igazi digitális bölcsészeti kihívás az, hogy a kutatói (történészi), technikai és könyvtári készségek, motivációk össze kell, hogy kapcsolódjanak. Ez nem csupán az ismeretek kibővítését jelenti, hanem az együttműködések egy magasabb szintjét is.

Igaza van Niels Brüggernek, hogy a kutatóknak nagyon sok segítségre van szüksége, hiszen ma még a történészek nem ismerik a webes archiválásban felmerülő, folyamatosan változó technikai kihívásokat, talán a webmegőrzés hosszú távú biztosításának jelentőségét sem.⁷⁶ Ezért a szakembereknek nem csupán a tulajdonképpeni feladatra, hanem arra is kell gondolniuk, hogy proaktív módon artikulálják az archiválás szükségességét. Mindezt annak érdekében, hogy ösztönözzék a meglévő kulturális örökségi intézményeket helyi, tematikus vagy nemzeti webarchívumok létrehozására, és megmutassák, szemléltessék e nagyszerű új forrásegyüttes felhasználásának lehetőségeit.

Web Archiving and Historical Research

Born digital content offers a more detailed and complete record of the present than what traditional sources provided of the past. The first section of this paper surveys current efforts and technologies to capture the present digital universe and reflects on the limitations of current web archives. The first attempts to archive the Hungarian web were made in 2017 in the National Széchényi Library, Budapest. The second part explores how this content could be harnessed for historical research, and how it will become the principal source of our recent past. The authors point out how web archives, and the resulting scale of data, will require new strategies and methodologies to deal with born digital sources effectively. They also show that born digital sources will also make it possible to pursue new types of inquiries that yield new results.

Keywords:

web archives, digital preservation, digital humanities, web historiography

⁷⁵ Vö. a témába vágó honlap, a *Web Archives for Historians* egyik konferenciabeszámolójával: Peter Webster, „What Do We Need to Know About the Archived Web?”, hozzáférés: 2018.05.22, <https://webarchivehistorians.org/2016/11/17/what-do-we-need-to-know-about-the-archived-web/>.

⁷⁶ Brügger, „Web History,” 316–325.

Markó Anita*ELTE Régi Magyar Irodalom Tanszék, Universität Wien Finno-Ugristik*

anita.marko.am@gmail.com

Hálózatok a 16–17. századi album amicorumokban: Az 1500 és 1700 közötti hungarika jellegű emlékkönyvbejegyzések hálózatelemzése az *Inscriptiones Alborum Amicorum* adatbázis alapján

Az *album amicorum*ok bejegyzései nem csak a filológiai elemzésre vagy egyes személyek biográfiai adatainak pontosítására adnak lehetőséget. A forráscsoport magában hordozza a hálózatként való interpretálás lehetőségét, a bejegyzések által összekapcsolódó hálózat vizsgálatát. Elemzésemben a szegedi működtetésű *Inscriptiones Alborum Amicorum* adatbázis anyagát használva az 1500 és 1700 közötti hungarika jellegű *inscriptiók* bejegyzőiből és albumtulajdonosaiból álló hálózat vizsgálatára teszek kísérletet. A hálózati modell adaptálását módszertani kísérletnek szánom. A legfőbb kérdésem az, hogy a hálózattudomány interdiszciplináris kerete és a társadalmi hálózatelemzés (matematikai és informatikai) eszközei miként használhatóak egy művelődéstörténeti korpuszra; feltárható-e a hálózati modell használatával a bejegyzők közötti, bejegyzések általi kapcsolódások mintázata, s hogy a megközelítésmód hozzájárulhat-e a már ismert adatok alapján új vagy eddig esetleg rejtőző jelenségek, filológiai, biográfiai részletek megvilágításához. Az elemzés során több mint 4000 bejegyzés adatait vizsgáltam, több időkeretben és bontásban. A vizsgálatban alkalmazott hálózatos elrendezés és elemzés rámutat olyan kiemelt értékű szereplőkre, elkülöníthető csoportokra, amelyek egy nagyobb, lineáris elrendezésű adathalmazban megbújhatnak. A hálózati modell pedig lehetőséget ad arra, hogy felfedhetővé váljanak gyenge kapcsolatok, ezáltal új kontextusban lássuk az *album amicorum*ok eddig ismert bejegyzéseit.

Kulcsszavak:

album amicorum, emlékkönyv, értelmiségi hálózatok, kora újkor, hálózatelemzés



Networks influence practically every subfield of literary studies.¹

1. A kiindulópont

A 16–18. században divatosnak számító emlékkönyvek, az *album amicorum*ok a legkézenfekvőbb korai példái a kapcsolatépítésnek, mai kifejezéssel élve a *networking*nek. Az eredetileg a peregrináló diákok által kedvelt, idézeteket és aláírásokat gyűjtő könyv, amely később a kor humanista barátságkultuszának műfajává vált szinte felkínálja a közösségi hálóként való interpretálás lehetőségét. A fennmaradt bejegyzések fontos forrásnak számítanak: nemcsak az idézetek filológiai elemzése vagy egyes személyek biográfiai adatainak pontosítása, de a tudás áramlásának mintázatai is kiolvashatók az adatokból. Az ilyen típusú áttekintő és összegző vizsgálatokra a 21. századi informatika eszközei új módszertani megközelítéseket adhatnak. A forráscsoport ugyanis nemcsak a nemzetközi, de a magyar kutatásban is jól feldolgozott: a szegedi működtetésű *Inscriptiones Alborum Amicorum* adatbázis² megfelelő kiindulópontot biztosít ahhoz, hogy átfogó és adatközpontú vizsgálatok készülhessenek. Elemzésemben az *album amicorum*okra tehát közösségi hálóként kívánok tekinteni, s az említett adatbázisban található 1500 és 1700 közötti *inscriptió*k bejegyzőiből és albumtulajdonosaiból álló hálózat vizsgálatára teszek kísérletet. A kor kulturális hálózatát összefűző sokrétű kapcsolategyüttesből tehát csupán egyet, a bejegyzések által megvalósult kapcsolatot kívánom kiemelni és fókuszba állítani. (A bejegyzések hálózatát az „irodalmi élet”, a kulturális mező egy nagyobb hálózatának részeként elemzem. Ezért megfelelő kísérleti példának tartom azt egy folyamatban lévő átfogó kutatáshoz, amely a magyar középkor és kora újkor irodalmi jelenségeit összekapcsoló értelmiségi hálózat feltárását tűzte ki célul.)

Ez a jelenlegi hálózatként való interpretálás egy módszertani kísérlet, amely során a legfőbb kérdés az, hogy a hálózattudomány interdiszciplináris kerete és a társadalmi hálózatelemzés (matematikai és informatikai) eszközei miként adaptálhatóak egy művelődéstörténeti korpuszra. Hipotézisem, hogy a hálózati modell használatával feltárható a bejegyzők közötti, bejegyzések általi kapcsolódások mintázata, s hogy a megközelítésmód hozzájárulhat a már ismert adatok alapján új vagy eddig esetleg rejtőző jelenségek, filológiai, biográfiai részletek megvilágításához.

¹ Patrick Jagoda, „Networks in Literature and Media,” in *Oxford Research Encyclopedia of Literature*, eds. Priscilla Wald and Paula Rabinowitz (Oxford: Oxford University Press, 2017), 1, hozzáférés: 2018.05.15, <https://doi.org/10.1093/acrefore/9780190201098.013.135>.

² Latzkovits Miklós, *Inscriptiones Alborum Amicorum*, 2003–2017, hozzáférés: 2018.05.15, <https://doi.org/10.14232/iaa>. Továbbiakban: IAA.

2. A kapcsolat

2.1. Az adatok forrása: *Inscriptiones Alborum Amicorum*

Az *album amicorum* „a 16. sz. közepétől a 18. sz. végéig élő könyvműfaj.”³ Az alapvetően kisméretű könyvekbe a diákok gyűjtötték peregrinációjuk során a bejegyzéseket: diáktársaiktól, professzoraiktól vagy patrónusaiktól. Később az arisztokrata ifjak körében is népszerű lett az emlékkönyv, amelybe az utazásaik közben meglátogatott főúri családok címereit gyűjtötték. Az eredetileg feltehetően wittenbergi környezetben kialakult, valószínűleg csak adminisztratív szerepet betöltő professzori bejegyzés így valójában annak mutatójává vált, hogy milyen kapcsolatépítő stratégiák kísérték az egyén életútját. Az albumok használata Melanchton nyomán kifejezetten divat lett; a bejegyzések, az albumversek külön műfajjá váltak. Az *inscriptiók* alapvetően három egységből álltak: egy mottóból (*symbolum*), amely a bejegyző többé-kevésbé állandó jelmondata; magából az általában idézetet tartalmazó bejegyzésből (*inscriptio*); a harmadik rész pedig az ajánlás (*dedicatio*), amelyben a bejegyző aláírása és a dátum is szerepel.⁴

A hungarika jellegű bejegyzések adatait az *Inscriptiones Alborum Amicorum* projekt publikálta adatbázisként. Az IAA meghatározása szerint hungarika jellegűnek minősülnek az alábbi bejegyzések:

- I. A magyarországi származású albumtulajdonosok emlékkönyveinek minden bejegyzése.
- II. A nem magyarországi származású albumtulajdonosok emlékkönyveinek magyarországi bejegyzőktől származó *inscriptiói*.
- III. Az ezen utóbbiakkal azonos oldalon olvasható egyéb bejegyzések.
- IV. A „grata vicinitas” esetei.
- V. A „Magyarországon” kelt bejegyzések.⁵

A bejegyzések adatait a készítőik kilenc blokkban rögzítették, tehát az adatbázis tartalmazza a bejegyzés keletkezésének időpontját, illetve az azzal kapcsolatos adatokat; a keletkezés helyét; a bejegyző jelmondatát; az esetleges ábrát és adatait; utólagos megjegyzéseket, amelyek lehetnek az albumtulajdonos jegyzetei vagy a *grata vicinitas* (szomszédossági) esetek, azaz

[...] azon esetek, amikor egy nem magyarországi származású albumtulajdonos emlékkönyvének valamely magyarországi bejegyzőtől származó *inscriptiója* valamilyen felirat pl. a „grata vicinitas”, „hae paginae jungunt amicos” – segítségével összefüggésbe van hozva egy szomszédos oldalon található, nem magyarországi bejegyzőtől származó *inscriptióval*.⁶

³ Földesi Ferenc, „Album amicorum,” (ID: 3096) in *Magyar művelődéstörténeti lexikon, VII. Középkor és kora újkor*, szerk. Kőszeghy Péter (Budapest: Balassi Kiadó, 2007), hozzáférés: 2018.05.15, http://mamul.btk.mta.hu/MAMUL_SZERK/mamul_view.php?editid1=3096.

⁴ Földesi Ferenc, „Album amicorum.”

⁵ „Az adatbázisról,” IAA, hozzáférés: 2017.06.22, <http://iaa.bibl.u-szeged.hu/index.php?page=about>.

⁶ „Az adatbázisról.”

Az adatbázis tartalmazza magát a bejegyzésszöveget; a dedikációval kapcsolatos és a bejegyző aláírásával kapcsolatos információkat. Mindehhez egy bibliográfiai blokk kapcsolódik, amely az adott albumhoz és a fent említett részletekhez ad szakirodalmi útmutatót. Az adatbázisból általam átvett (és külső felhasználóként átvehető) adatok a bejegyzések azonosító számát, a bejegyző nevét, az albumtulajdonos nevét, a bejegyzés idejét, helyét, a bejegyzést tartalmazó forrás megjelölését tartalmazzák.

Az adatbázis szerinti legelső hungarika jellegű bejegyzés 1553. június 22-én keletkezett, német nyelven, bejegyzője a Budáról származó Christopher Mandel. Az adatbázis jelen pillanatban utolsó bejegyzése 1826-ból való. Az adatbázis folyamatosan bővül, ezért fontosnak tartom kiemelni, hogy a vizsgálat elkészüléséhez szükséges adatok az IAA 2017. május 29-i állapotát tükrözik. Az adatbázist készítő kutatócsoport a legnagyobb releváns könyvtárakkal, levéltárakkal, illetve nemzetközi intézményekkel és magánszemélyekkel is együttműködik. A forrásanyag feldolgozottságáról Latzkovits Miklós, a kutatócsoport vezetője 2015-ben elmondta, hogy addig 300 albumból gyűjtötték ki a bejegyzéseket, amely összesen 11991 darabot jelent.

108 album tulajdonosa volt magyarországi (vagy erdélyi) személy, ezeket tehát (néhány jelenleg is munka alatt lévő kivételtől eltekintve) már teljes egészében átírtuk az IAA-ban, ami 11.249 bejegyzést jelent. Utóbbiak kapcsán bejegyzés átlagot is tudunk mondani. Eddigi tapasztalataink szerint átlagosan 107 bejegyzéssel lehet számolni a magyarországi személyek által vezetett emlékkönyvek esetében. Az ő albumaikból előkerült bejegyzések a teljes anyag mintegy 94 százalékát teszik ki. Az általunk épített, érzésünk szerint is nagyon töredékes „albumkatalógusban” jelenleg 304 ma is elérhető olyan 16–18. századi albumot bibliografizáltunk, melynek tulajdonosa magyarországi személy volt. Ez tehát nagyjából 32.000 bejegyzést jelent.⁷

Elmondása szerint 2017-re a regisztrált bejegyzések száma 12500-ra nőtt, az említett albumkatalógus is bővült. Óvatos becslése szerint az anyag felét-harmadát ismerhetjük és arra következtet, hogy ma mintegy 60000–100000 hungarika jellegű bejegyzés lehetne elérhető a 16–18. századból. Mivel azonban sok album magántulajdonban van, arról nehezen szerezhető információ, hogy hány olyan hungarika jellegű bejegyzés létezik jelenleg, amiről egyáltalán nem tudunk. Épp ezért azt is nehéz volna megmondani, hogy mennyi hungarika vonatkozású albumbejegyzés keletkezett a 16–18. században. A nagyságrend megbecsüléséhez Latzkovits az időszakban a német egyetemeken tanuló magyarországi diákok számát említi. Körülbelül 27000 fővel számol, s feltételezi, hogy „inkább mindegyiküknek, mint csaknem mindegyiküknek” volt albumuk:

Az eddigi munka alapján kiszámoltuk, hogy egy magyarországi tulajdonos albumában átlagosan (és meglepő módon) 100 bejegyzés található. Márpedig ez nagyjából 2.700.000 valaha megírt albumbejegyzést jelent csak a németországi egyetemek hallgatóinak mára már jobbára megsemmisült emlékkönyveiben.

⁷ Az adatbázis jelenlegi (2017.01.07.) állapotáról, a forrásanyag feldolgozottságának mértékéről, és a becslésekről szóló adatokat Latzkovits Miklós személyes közlése alapján idézem, a kapott információt ezúton is köszönöm neki (2017.01.07.).

Az összes létező bejegyzésnek tehát csupán töredéke maradhatott fenn, amelynek feldolgozottsága vélhetően a 30–50%-ot éri el. Ez azt jelenti, hogy új bejegyzések, új emlékkönyvek folyamatosan kerülnek, illetve kerülhetnek elő, majd építhetők be az adatbázisba. A hálózati modell alkalmazása azonban ebből a szempontból jól alkalmazkodik: a modell bővíthető, frissíthető; új kapcsolatok és új elemek bármikor hozzáadhatók, a végzett számítások aktualizálhatók. Az adatok hiányossága azonban nem problémamentes, hiszen épp emiatt általános, a teljes, valaha létező korpuszra nem szerencsés erős következtetéseket levonni.

2.2. Az adatok kiválasztásának szempontjai

A fentiekből következik tehát, hogy a feldolgozott adatok egy vetületét adják az *album amicorum*-használatnak. Érdekes ezért véletlen mintavételként tekinteni és használni a feldolgozott és elérhető adatokat, mely esetben a véletlenszerű leválogatást a forrásmegmaradás és az adatfeldolgozás esetlegessége hozta létre. A mintavételi eljárás nem reprezentatív: ez azonban nem jelenti azt, hogy nem lehet releváns eredményekre jutni. Figyelni kell viszont arra, hogy eredmények csak a konkrét mintára vonatkozathatók, és nem szabad azokat a teljes populációra általánosítani. Más a kimutatható mintázatok referencialitása: elsősorban arra lehet következtetni, hogy az általunk jelenleg ismert bejegyzők és bejegyzőik milyen viszonyban álltak egymással; arról, hogy a mintavételben működő folyamatok mint hálózatkötő tényezők a hálózat nem ismert részeiben is megjelennek, csak valószínűségi feltételezéseket lehet tenni; a már ismert kapcsolatok és viszonyok azonban indikátorként működnek.

A jelenlegi munkának nem célja az *album amicorum*ok teljes nemzetközi hálózatának vizsgálata, s a magyar vonatkozású kapcsolatoknak ezek tágabb kontextusában való elemzése – bár ez szintén releváns és izgalmas kutatási irány lenne. Ahhoz azonban, hogy akár egy rövidebb időszakra nézve is kiegészítsük az adatbázist például a nem magyar albumtulajdonosok nem magyar bejegyzők általi bejegyzéseivel (ezeket nem gyűjti az IAA), csupán akkor lenne lehetőség, ha bevonnánk a kizárólag hungarika anyagot feldolgozó adatbázison kívül egyéb nemzetközi adatbázisok eredményeit, illetve amennyiben nem készült ilyen, akkor egyéb (levéltári) forrásanyagokat. Ez a kutatás meghaladná jelen tanulmány kereteit.

A vizsgált hálózat másik leszűkítése időbeli: az általam kijelölt időszak 1500 és 1700 közötti, tehát a legelső 1553-as bejegyzéstől az 1692 és 1740 között keletkezett bejegyzéscsoportig tart. A szigorúan időponthoz kötött adatfelvétel módszere bizonyos szempontból nem adhat teljes képet: azoknak az albumtulajdonosoknak a bejegyzői hálózata nem látható egészében, amelyek 1700 után keletkezett bejegyzéseket is tartalmaznak. A jelenlegi vizsgálat szempontjából azért tettem meg ezt a leválasztást, mert a vizsgálatban csak azokat a kapcsolatokat kívántam elemezni, amelyek 1700 előttiak. Emellett az 1553-tól 1700-ig terjedő adatokból kirajzolódó hálózat referenciaként és összehasonlításként áll – a hálózat növekedésének nagyságrendje megállapítható, mivel azonban a 150–200 év túl nagy időtáv a szociális kapcsolatok értelmezéséhez, kiemelek egy időben (és számosságban is) kisebb hálózatot az elemzés számára. Ebbe az első 1553-as bejegyzéstől kezdve beletartoznak azok az albumok is, amelyek legalább

egy darab 1600 előtt keletkezett bejegyzést tartalmaznak. A mintavétel konkrétumait a 4.3 alfejezet tartalmazza.

Az elemzés elvégezhető lenne úgy is, ha viszonyítási pontként és a vizsgálat alapjául nem a bejegyzést, hanem a forrásművet választanánk. Ez esetben az adatokat úgy kellene válogatni, hogy azok az albumok már ne kerüljenek bele a hálózatba, amelyek 1700. január 1. utáni *inscripti*ót alkalmaznak. Jelen esetben azonban fontosabbnak tartottam a kapcsolatok számontartását, mint az albumokét. Az egyes emlékkönyvek hálózatainak elemzése másik vizsgálat lenne, amely egy hibrid hálózatként, latouri értelemben vett cselekvőhálózatként lenne interpretálható.⁸

Jelen esetben a hálózatot szociológiai megközelítésben értelmezem, tehát személyközi kapcsolatokból álló rendszernek tekintem, amely gráfként vizualizálva csúcsokból (pontok, aktorok vagy cselekvők) és élekből (linkek, kapcsolatok) áll. A hálózat pontjait a bejegyzők és az albumtulajdonosok, azaz a bejegyzések címzettjei jelentik, a közöttük fennálló kapcsolat pedig maga a bejegyzés.

3. A hálózat

Reménytelenül bonyolult rendszerekkel vagyunk körülvéve.⁹

3.1. A hálózattudományi szemüveg

Az utóbbi évtizedekben a hálózat és a háló fogalmát több tudományterület (a biológiától kezdve a társadalomtudományon keresztül az informatikáig) is meghatározó interpretációs keretként és leírási modellként kezdte használni. Ez a transzdiszciplináris alkalmazás vezetett oda, hogy a hálózatok vizsgálata az elmúlt két évtizedben ugrásszerűen fejlődött, és új tudományággá vált. A hálózat kutatás (*network science*) fő felfedezésének éppen azt tartja, hogy „a tudomány, a természet és a technológia különböző területein előforduló hálózatok szerkezete igen hasonló, mert ezeket ugyanazok a rendezőelvek alakítják. Emiatt ezeket a hálózatokat ugyanazokkal a matematikai eszközökkel kutathatjuk.”¹⁰

Mindezek mellett meg kell említeni, hogy a fogalmak használata éppen emiatt a transzdiszciplinaritás miatt nem konzekvens: a hálózattudomány és a hálózat kutatás szinonimaként jelenik meg, de a hálózatelmélet és a hálózatelemzés (*SNA*, *social network analysis*) definícióját is gyakran emlegetik. Utóbbi kettőt alapvetően a szociológia használta eddig: a társadalmi rendszerek működésének elméleti és módszertani keretének megjelölésére. Az általam használt elemzési módszereket és definíciókat főként a szociális hálózatelemzés módszertani készletéből merítettem, de mivel az elemzések alapját adó matematikai eszközök megegyeznek a hálózattudomány elemzési

⁸ Bruno Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory* (New York: Oxford University Press, 2005). A cselekvőhálózat-elmélet (*actor-network-theory*) olyan hálózatként értelmezi a jelenségeket, amelyben azonos értékű cselekvőnek – *aktor* helyett *aktant* – tekint mind az élő személyeket, mind a tárgyi világ elemeit.

⁹ Barabási Albert-László, *A hálózatok tudománya* (Budapest: Libri, 2016), 39.

¹⁰ Barabási, *A hálózatok tudománya*, 42.

módszereivel, így a tanulmányban használt módszertani keretet a hálózattudomány társadalmi rendszerekre specializált részének tekintem.

A hálózattudományi vizsgálatok célja megérteni a komplex rendszerek univerzális tulajdonságait és eredetüket, felfedni az ezen hálózatok fejlődését alakító törvényeket, és azt, hogy ezek a törvényszerűségek miként vannak hatással a hálózat viselkedésére.¹¹ A tudományág kialakulását és eszköztárát tekintve is interdiszciplináris: különböző területeken ugyanis már évtizedek óta egymástól függetlenül is elkezdődött a hálózatok vizsgálata. A legkorábbi megközelítést a matematikai gráfelmélet nyújtja: Leonard Euler 1736-os königsbergi hidakról szóló példája máig alapmintának számít. A társadalmi kapcsolatháló vizsgálat az 1970-es években kezdődött, az első kutatások Mark Granovetter nevéhez fűződnek. A komplex rendszerek összefüggéseinek feldolgozásához elengedhetetlenül szükségesek a számítástechnikai eszközök, így az informatika fejlődése rendkívüli ugrást jelentett az ilyen irányú vizsgálatok számára. Összegezve tehát azt lehet mondani, hogy a hálózatkutatás módszertana alapvetően empirikus, gráfelméleti-matematikai, kvantitatív és informatikai.¹²

A legfontosabb univerzális tulajdonságok a *kisvilágság*, a *skálafüggetlenség*, az egymásba-ágyazottság és a *gyengekapcsoltság*. A *kisvilágság* jelenségének első példáját a nemzetközi szakirodalom is Karinthy Frigyes nevéhez köti. 1929-ben, *Láncszemek* című novellájában¹³ veti fel azt a kérdést, hogy „hány ismeretségi kapcsolaton keresztül lesz összeköthető két véletlen személy”. Ezt később tudományos keretek között a szociológia vizsgálta: Stanley Milgram harvardi kísérleti pszichológus a személyközi kapcsolatok egyes embereket közösséggé szervező hálóját kutatva fedezte fel az úgy nevezett „hatlépésnyi távolságot.”¹⁴ A vizsgálat eredménye szerint két egymást sohasem ismerő, tökéletesen véletlenszerűen kiválasztott személy közt egy levél célbejuttatásához átlagosan 5,5 ember szükséges. A következtetés tehát, hogy világunk szorosan összekapcsolt, „bárki csak néhány kézfogásra van bárki másától.”¹⁵

A hálózatok egyik alapmodelljeként sokáig a véletlen gráfokat leíró Erdős–Rényi-modellt tartották számon. Ebben az elméletben központi az egyenlőség gondolata: majdnem minden ponthoz azonos számú él tartozhat.¹⁶ A valóságban létrejövő hálózatok azonban nem statikusak, s emiatt más modell érvényes rájuk – erre hívta fel a figyelmet Barabási Albert-László és Albert Réka kutatása.¹⁷ Modelljük alapvetése, hogy a hálózatok növekedésének szervezője a preferenciális kapcsolódás – tehát az új pontok nem véletlenszerűen kapcsolódnak a hálózat egy pontjához, hanem a legnagyobb fokszámú pont felé gravitálnak.

¹¹ Barabási, *A hálózatok tudománya*, 42.

¹² Barabási, *A hálózatok tudománya*, 43–44.

¹³ Karinthy Frigyes, „Láncszemek,” in Karinthy Frigyes, *Címszavak a Nagy Enciklopédiához: Cikkek* (Budapest: Szépirodalmi Könyvkiadó, 1980), Magyar Elektronikus Könyvtár, <http://mek.oszk.hu/07300/07367/html/01.htm#54>.

¹⁴ Stanley Milgram, „The small world problem,” *Psychology Today* 1, 1. sz. (1967): 61–67.

¹⁵ Barabási Albert-László, *Behálózva* (Budapest: Magyar Könyvklub, 2003), 28.

¹⁶ Csermely Péter, *A rejtett hálózatok ereje: Hogyan stabilizálják a világot a gyenge kapcsolatok?* (Budapest: Vince Kiadó, 2004), 10; Péter Erdős és Alfréd Rényi, „On Random Graphs, I,” *Publicationes Mathematicae* 10 (1959): 290–297.

¹⁷ Albert-László Barabási and Réka Albert, „Emergence of Scaling in Random Networks,” *Science* 286, 5439. sz. (1999): 509–512, <http://doi.org/10.1126/science.286.5439.509>.

A preferenciális kapcsolódás egy valószínűségi mechanizmus: egy új csomópont szabadon kapcsolódhat a hálózat bármelyik csomópontjához, legyen az egy középpont vagy olyan, amelynek eddig csak egy kapcsolata volt. [...] ha az új csomópont egy 2 vagy egy 4 fokszámú csomóponthoz is kapcsolódhat, akkor kétszerre nagyobb eséllyel választja a 4 fokszámú csomópontot.¹⁸

Ez az elrendeződés hatással van az adott hálózat szerkezetére is: kedvez a diverzitásnak, segíti a hálózaton belüli tájékozódást, a hálózatot ellenállóbbá teszi. Kialakulnak nagy fokszámú, azaz sok kapcsolattal rendelkező összekötők, amelyek kiemelkednek, láthatóak lesznek a többi tag számára, s ezáltal a folyamatok elindítói, trendek irányítói lehetnek.¹⁹

A gyenge kapcsolatok fogalma²⁰ ugyanígy fontos a kulturális hálózatok értelmezésében. Ezek kis valószínűségű és rövid idejű kölcsönhatások, de a hálózatok legfontosabb pontjain találhatóak – így biztosítják a stabilitást, de nem merevítik meg a hálózatot.²¹ A Berlow-féle funkcionális definíció szerint „egy kapcsolatot akkor nevezünk gyengének, ha hozzáadása vagy elvétele nem befolyásolja statisztikailag kimutatható mértékben a hálózat külső paramétereinek átlagértékét”. A különböző levezetések szerint nagyjából a hálózatoknak csak 20%-a erős kapcsolat.²² A gyenge kapcsolat a gyakorlatban azt jelenti, hogy például álláskeresésnél egy véletlen beszélgetés (utazás közben, repülőn, és így tovább) előbb juttathat munkához minket, mint egy közeli barátunk vagy családtagunk. Hiszen *erős kapcsolataink* ugyanazokat a személyeket ismerik, mint mi; de egy távoli ismerős, egy beszélgetőtárs hozzákapcsol minket egy számunkra addig ismeretlen körhöz, amely új lehetőségeket ígér. Tehát a gyenge kapcsolat nagyobb valószínűséggel hoz új információt, tartalmat a hálózatba, mint az erős kapcsolatok. Ez a jelenség – ahogy a fent említettek is – a hálózattudomány alapvetése szerint éppúgy igazak arra a hálózatra, amelynek pontjait személyek, összekötő éleit pedig a köztük lévő kapcsolatok alkotják.

3.2. Mérőszámok és definíciók

A jelen vizsgálat szempontjából ez azt jelenti, hogy az általam kiválasztott csoportban – az egymást bejegyzéseik által ismerősként megjelölő albumhasználók körében – megkísérelhetjük a kapcsolatok sűrűségének, a személyek közt kialakult hierarchiának és presztízsnak, a klikkek és klaszterek kialakulásának vizsgálatát, a hálózatban működő összekötők, hidak, gyenge kapcsolatok szerepének felmutatását és elemzését. Az elemzésnek a hálózat társadalmi hálózatelemzésben használt definícióját használom, miszerint:

¹⁸ Barabási, *A hálózatok tudománya*, 188.

¹⁹ Barabási, *Behálózva*, 35.

²⁰ Mark Granovetter, „The Strength of Weak Ties: A Network Theory Revisited,” in *Social Structure and Network Analysis*, eds. P. V. Marsden and N. Lin (Beverly Hills, CA: Sage Publications, 1982), 105–130. Magyar fordításban: „A gyenge kötések ereje: A hálózatelemzés felülvizsgálata,” in *Társadalmak rejtett hálózata*, szerk. Angelusz Róbert és Tardos Róbert (Budapest: Magyar Közvéleménykutató Intézet, 1991), 371–400.

²¹ Csermely, *A rejtett*, 85.

²² Berlow-definíció: Csermely, *A rejtett*, 85.

[...] a véges számú pontot teljesen vagy részben összekötő vonalak halmazát nevezzük hálónak. A viszonyok egy bizonyos típus által meghatározott pontok halmazát nevezzük a továbbiakban hálózatnak. Ha a vizsgált pontok legalább egy részhalmazát többféle viszony szerint rendezzük, akkor a továbbiakban többretegű hálózatokról (multiplex networks) fogunk beszélni.²³

A hálózatok elemzésének²⁴ alapfogalmai a (csomó)pontok (csúcsok, aktorok, cselekvők) és a köztük lévő kapcsolatok (élek, linkek). Egy hálózat elsődlegesen meghatározandó tulajdonságának számít, hogy az őt alkotó kapcsolatok irányítottak vagy irányítatlanok. Irányított kapcsolat esetén aszimmetrikus hálózat jön létre, mint az *inscriptiók* esetében is. A kapcsolatok iránya a bejegyzők felől a tulajdonosok felé mutat. (Ha *A* írt bejegyzést *B*-nek, az nem jelenti, hogy *B* is írt *A*-nak, a hálózat szempontjából ez azért jelentőségteljes, mert az információk és egyéb a hálózatban áramló értékek közlekedésének iránya meghatározott.)

Míg maga a bejegyzés irányított kapcsolatot hoz létre, addig ennek az összeköttetésnek a tartalma, miszerint a bejegyző ismeri a személyt, akinek az albumába ír és fordítva, szimmetrikus kapcsolatként is értelmezhető és elemezhető. A legalapvetőbb mérőszámoknak a hálózatot alkotó pontok számát (*n*) és a hálózatban létrejött összes kapcsolat számát (*L*) tekinthetjük. A pontok legfontosabb tulajdonsága a *fokszám* (*degree*, *k*), amely az adott ponthoz tartozó kapcsolatok számát jelöli. Irányított esetben ez kétféle lehet: a *kimenő fokszám* (*outdegree*, „ki-fok”) azon kapcsolatokat méri, amelyek a cselekvő felől indulnak más cselekvők irányába; a *bemenő fokszám* (*indegree*, „be-fok”) ennek épp ellenkezője: a pont felé irányuló kapcsolatok mérőszáma. Fontos megjegyezni, hogy egy hálózatban a távolság fogalma nem érzékelhető távolságot jelent – a vizualizált gráfban nincs tartalmi szerepe az élek hosszának. A két cselekvő közti *távolságot*²⁵ (*distance*, *d*) a köztük lévő *út* (*path*) *hossza*, azaz az utat alkotó kapcsolatok száma jelenti. Egy hálózat *átmérője* (*diameter*) pedig a legrövidebb úthosszak maximuma. (A mérőszámot csak összefüggő hálózatokra lehet alkalmazni, definíció szerint ugyanis a nem összefüggő gráfok átmérője végtelen.)

A hálózat értelmezésében fontos szerepet játszik az összekapcsoltság vizsgálata, a hálózat összefüggő részhálózatainak (*komponens*)²⁶ meghatározása. Irányítatlan kapcsolatok esetén az a hálózat számít összefüggőnek, amelyben minden pont elérhető a

²³ Takács Károly, szerk., *Társadalmi kapcsolathálózatok elemzése* (Budapest: BCE Szociológiai és Társadalompolitikai Intézet, 2011), 5, https://www.tankonyvtar.hu/hu/tartalom/tamop425/0010_2A_08_Kapcsolathalo_elemzes_szerk_Takacs_Karoly/0010_2A_08_Kapcsolathalo_elemzes_szerk_Takacs_Karoly.pdf.

²⁴ A hálózatelemzés alapfogalmaihoz és módszertanához lásd: Barabási, *A hálózatok tudománya*; Christian Stegbauer und Roger Häußling, Hg., *Handbuch Netzwerkforschung* (Wiesbaden: Springer VS, 2010); Takács, *Társadalmi*.

²⁵ „A távolság számításához ismernünk kell a séta (*walk*), a vonal (*trail*) és az út (*path*) fogalmakat. A séta pontok és élek olyan sorozata, mely ponttal kezdődik és azzal is végződik, egy pontot mindig hozzá illeszkedő él előz meg és az is követ a sorozatban, a pontok és élek többször is előfordulhatnak. A vonal olyan séta, melyben az élek nem ismétlődnek a sorozatban, az út esetén pedig a pontok sem fordulhatnak elő egynél többször (ilyenkor az élek sem ismétlődhetnek). A séta, a vonal és az út hossza minden esetben a benne szereplő élek száma. Két pont közti távolság a két pont közötti legrövidebb út hosszával egyenlő. Ha két pont közt nincs út, a távolságot végtelennek definiáljuk.” Takács, *Társadalmi*, 24.

²⁶ Barabási, *A hálózatok tudománya*, 79.

hálózat egy másik pontjából. Irányított kapcsolat esetén megkülönböztetünk erősen és gyengén összefüggő hálózatot: előbbiben oda-vissza kapcsolatokon keresztül érhető el minden pont egymástól, utóbbiban csak egyirányú kapcsolatok tartják a hálózatot. Ha két komponenst egyetlen él köt össze, a kapcsolatot *hídnak* nevezzük: ennek elvágásával a hálózat megszűnik összefüggőnek lenni.

Ezen felül fontos a hálózat *sűrűsége* (*density*, Δ), amely nemcsak a szomszédos kapcsolódásokat, hanem az összes lehetséges és a ténylegesen megvalósult kapcsolat arányát méri;²⁷ a *központiság* (*centrality*) pedig azt, hogy a létező kapcsolatok közül mennyit birtokol az adott szereplő.

A *presztízst* (*prestige*) irányított kapcsolatok esetén tudjuk kifejezni. Fontos megjegyezni azonban, hogy itt nem a köznapi tekintély, hatalom értelemben kell a szót értelmezni: jelentése szerint annak a szereplőnek nagy a *presztízse*, amelynek nagyobb a bemenő fokszáma, mint a kimenő; míg a *befolyás* ennek fordítottja: a nagyobb kimenő fokszámot mutatja. A hálózat szimmetrikus kapcsolatainak fontos mutatói a már említett *központiságértékek* – ezek azt mutatják meg, hogy mennyire fontos az adott szereplő. Ez azonban nem egzakt fogalom, a fontos szereplők a hálózat stratégiai pontjain helyezkednek el, de ez a fontosság többféle módon értelmezhető. Központi személynek tekinthetjük azt, akinek a legnagyobb a kapcsolati aktivitása, akihez sokan kapcsolódnak, vagy aki minél több szereplővel tart fenn minél szorosabb kapcsolatot. Olyan szereplőket is kiemelhetünk továbbá, akik hálózatmegszakító pozícióban vannak.²⁸ A legegyszerűbb az előbb említett fokszámok alapján vett *fok-centralitást* nézni (*degree centrality*, *CD*),²⁹ eszerint minél több kapcsolattal rendelkezik egy adott szereplő, annál központibb szerepet tölt be a közösségben. Egy aktor *közelségi központisága* (*closeness centrality*, *CC*)³⁰ azt jelenti, hogy a legtöbb szereplőt közvetlenül vagy kevés lépéssel el tudja érni, míg más szereplők nem. A *közöttiségcentralitás* (*betweenness centrality*, *CB*)³¹ azon alapul, hogy a szereplő hálózatbeli sikere attól függ, hány másik pont vagy csoport közt tölt be közvetítő szerepet.

Míg a centralitásértékek az irányítatlan kapcsolatokat jellemzik, addig ezeknek a mérőszámoknak az irányított kapcsolatokra való megfeleltetését jelentik a *presztízsmutatók*. A *fok-presztíz* (*degree prestige*, *PD*) megegyezik az adott pont bemenő fokszámával, tehát azt a szereplőt emelhetjük ki vele, akit sokan választanak. A másik lehetőség a *szomszédsági presztíz* (*proximity prestige*, *PP*) kiszámítása, amely a közelségcentralitáshoz hasonlóan azt méri, hogy a szereplők milyen közel vannak a kiválasztott aktorhoz. Ez a távolság azért nem egyezik meg feltétlenül az irányítatlan hálózatával, mert két azonos pont közt a kapcsolatok irányítottsága miatt oda-vissza más úthosszak lehetnek jellemzőek. A harmadik fontos mutató a *rangpresztíz* (*rank prestige*, *PageRank*, *PR*), amelynek lényege, hogy azt is figyelembe vesszük, milyen tulajdonsággal rendelkeznek azok, akik az adott szereplőt választják: tehát, akit egyébként

²⁷ Irányítatlan hálózatban: $\Delta = \frac{\text{létrejöttkapcsolatokszáma}}{\text{összeslehetségeskapcsolatszám}} = \frac{L}{\frac{n(n-1)}{2}}$; irányított hálózatban: $\Delta = \frac{L}{n(n-1)}$.

²⁸ Takács, *Társadalmi*, 22.

²⁹ Adott csomópont (i) fok-centralitása megegyezik a fokszámával: $C_{D_i} = k_i$.

³⁰ Adott pont (i) közelségcentralitása fordítottan arányos a közte és a többi szereplő közti távolságok összegével: $C_{C_i} = \frac{1}{\text{távolságokösszege}}$.

³¹ Azoknak az utaknak az összege, amelyek minimális hosszúságúak, és keresztülhaladnak az adott szereplőn: $C_{B_i} = \frac{\text{ipontonáthaladóutakszáma}}{\text{összesútszáma}}$.

marginális szereplők választanak, az kisebb értékkel rendelkezik, mint akit magas presztízsértékű tagok választanak.³²

Ezek az adatok és mérőszámok kis elemszámú hálózat esetén a gráfként való ábrázolás során akár szemmel láthatóak, de nagyobb elem- és kapcsolatszám esetén a grafikai kép átláthatatlan, ezért mátrixalapú elemzésekre és a matematikai-statisztikai alapú eredményekre lehet alapozni.³³

4. Az irodalom hálójában

A társadalom hálózatként való interpretációja már a '70-es években elkezdődött, a hálózat metaforaként és interpretációs lehetőségként az irodalomtudományban elsősorban a 21. században vált fontossá, összefüggésben az új médiumok megjelenésével. Emellett azonban a történeti és irodalomtörténeti kutatások egyre inkább felhívják a figyelmet arra, hogy a biográfiai adatok vagy a művelődéstörténet eseményei és szereplői nem önmagukban álló figurák: az identitás, a társadalmi szerepek és életpálya-modellek mind a kontextus és a szereplőket magukba foglaló hálózatok teremtményei.

A hálózatok esztétikájával foglalkozó Patrick Jagoda foglalja össze, milyen megközelítésben alkalmazzák a kultúratudományok a hálózatokat. Nem teljesen elkülöníthető kategóriákról beszél, de nagyjából négy területet említ meg: a történeti kutatásokat,³⁴ a digitális bölcsészet körébe tartozókat, az irodalomszociológiai és a formalista irányt.³⁵ A hálózatok alkalmazhatóságának sokféleségét azzal magyarázza, hogy a hálózatokban megjelenő komplexitás és összekapcsoltság egybecseng az intertextualitás és a szoros olvasás tapasztalatával.³⁶ Az egyes vizsgálatok azonban nem egyértelműen különíthetők el a fenti felosztás szerint: nem módszerükben, hanem fókuszpontjukban különböznek. Míg Moretti az úgynevezett *távoli olvasás* (*distant reading*) megközelítésmódjával a szövegen belüli figurákat és cselekményeket elemzi (legismertebb a *Hamlet*-elemzése), addig a Jagoda által irodalomszociológiainak mondott vizsgálat – ugyanazokkal a digitális bölcsészeti eszközökkel – az irodalom intézményrendszerének és társadalmi beágyazottságának aktoraira és eseményeire koncentrálnak. Példaként említhető Natalie M. Houston projektje,³⁷ amely a viktoriánus költészet 1840–1900 között működő kapcsolathálóját térképezi fel; illetve Amy Hungerford 2016-ban megjelent könyve, mely az olvasók hálózatát értelmezi a könyvkiadás, a szerzők és

³² „Mivel ez minden kapcsolatháló-szereplőre igaz, a rangok láncolatát kell feltérképeznünk egy adott szereplő rangjának meghatározásához. Ezt legegyszerűbben mátrixalgebrai úton tehetjük meg, sajátértékek számításával.” Takács, *Társadalmi*, 26.

³³ Az elemzéseket a *Microsoft Excel 2010* adatbázis-kezelő programmal és a *Gephi* (v. 0.9.2) nyílt forráskódú hálózatelemző és vizualizáló szoftverrel végeztem.

³⁴ Jonathan H. Grossman, *Charles Dickens's Networks: Public Transport and the Novel* (Oxford: Oxford University Press, 2012); Alexander Galloway, „Networks,” in *Critical Terms for Media Studies*, eds. W. J. T. Mitchell and Mark B. N. Hansen (Chicago: University of Chicago Press, 2010), 280–296; Caroline Goodson, Anne E. Lester, and Carol Symes, *Cities, Texts and Social Networks, 400–1500: Experiences and Perceptions of Medieval Urban Space* (Farnham: Ashgate, 2010).

³⁵ Caroline Levine, *Forms: Whole, Rhythm, Hierarchy, Network* (Princeton: Princeton University Press, 2015).

³⁶ Jagoda, „Networks,” 1, 4.

³⁷ Natalie M. Houston, „Toward a Computational Analysis of Victorian Poetics,” *Victorian Studies* 56, 3. sz. (2014): 498–510, <http://doi.org/10.2979/victorianstudies.56.3.498>.

a könyvterjesztők kontextusában.³⁸ Jagoda végkövetkeztetésével egyetértve az ilyen típusú vizsgálatokban a hálózatban való modellezés nemcsak vizuális segédeszköznek számíthat, hanem felfedhetővé teheti az irodalom közlekedését, recepcióját, működését a társadalom különböző rendszereiben.³⁹

A magyar irodalomtörténeti kutatásban legutóbb Bátor Anna⁴⁰ és Antal Alexandra⁴¹ foglalkozik hálózati modellben történő interpretációval. Míg Antal irodalomszociológiai kontextusban a *Bécsi Magyar Hírmondó* szerkesztőségi hálózatát elemzi, addig Bátor Anna filológiai megközelítésben használja a hálózat interpretációs keretét, és tanulmányában Wallaszky Pál *Conspectus* című művének kompilációs gyakorlatát vizsgálja.

A jelenleg futó nemzetközi hálózattudományi alapú kutatások közül két példát említek: az első az inkább történeti irányú *Six Degrees of Francis Bacon*⁴² elnevezésű projekt, amely a filozófus körüli *énhálózat* feltárására törekszik. Ezáltal egy multiplex, több kapcsolattípust felvonultató közösséget tesz megközelíthetővé. A *Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic*⁴³ című holland kutatás a korszak leveleinek adatbázisban történő rögzítését, digitális publikálását és a levelezők kapcsolathálójának megalkotását célozza. A projekt nemzetközi együttműködést tart fenn több hasonló kutatással: az Oxfordi Egyetem *Cultures of Knowledge: An Intellectual Geography of the Seventeenth-Century Republic of Letters*⁴⁴ és a Stanford Egyetem *Mapping the Republic of Letters*⁴⁵ munkacsoportjával. Az itt vizsgált hálózat szereplőit egyféle kapcsolat, a levelezés köti össze, ebben működésük nagyon hasonlít az *album amicorum*ok adatainak feldolgozhatóságához.

5. Közösségi háló – az 1500–1700 közti bejegyzések és bejegyzők hálózatai

Az *Inscriptiones Alborum Amicorum* adatbázis elérhető adatai közül a hálózatok vizsgálatához az alábbi adatcsoportokat dolgoztam fel: a bejegyzés rekordszáma (*ID*), a bejegyző (*source*), a bejegyzés címzettje (*target*), a bejegyzés albumának jelzete (*sign*), a bejegyzés dátuma (*date*), a bejegyzés időköre (*yeartype*), a bejegyzés helye (*place*). Jelenleg csak azokat a legalapvetőbb adatokat vontam be a vizsgálatba, amelyek az adatbázis online felületéről külső felhasználóként tölthetők le.

³⁸ Amy Hungerford, *Making Literature Now* (Stanford, CA: Stanford University Press, 2016).

³⁹ Jagoda, „Networks,” 7–8.

⁴⁰ Bátor Anna, „A tudás hálózatai: Wallaszky Pál historia litterariája és a 18. századi tudástranszfer,” *Irodalomismeret* 27, 3. sz. (2016): 35–63.

⁴¹ Antal Alexandra, „A bécsi Magyar Hírmondó (1789–1803) szerkesztői hálózata,” *Irodalomtörténeti Közlemények* 118, 1. sz. (2014): 99–117.

⁴² Christopher Warren, Daniel Shore, Jessica Otis, Lawrence Wang, Mike Finegold and Cosma Shalizi, *Six Degrees of Francis Bacon*, Carnegie Mellon University, hozzáférés: 2018.05.15, <http://www.sixdegreesoffrancisbacon.com/>.

⁴³ Wijnand Mijnhardt, *Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic*, v. 1.0, hozzáférés: 2018.05.15, <http://ckcc.huygens.knaw.nl/epistolarium>.

⁴⁴ Howard Hotson, *Networking the Republic of Letters, 1550–1750*, University of Oxford, hozzáférés: 2018.05.15, <http://www.culturesofknowledge.org/>.

⁴⁵ Dan Edelstein and Paula Findlen, *Mapping the Republic of Letters*, Stanford University, hozzáférés: 2018.05.15, <http://republicofletters.stanford.edu/>.

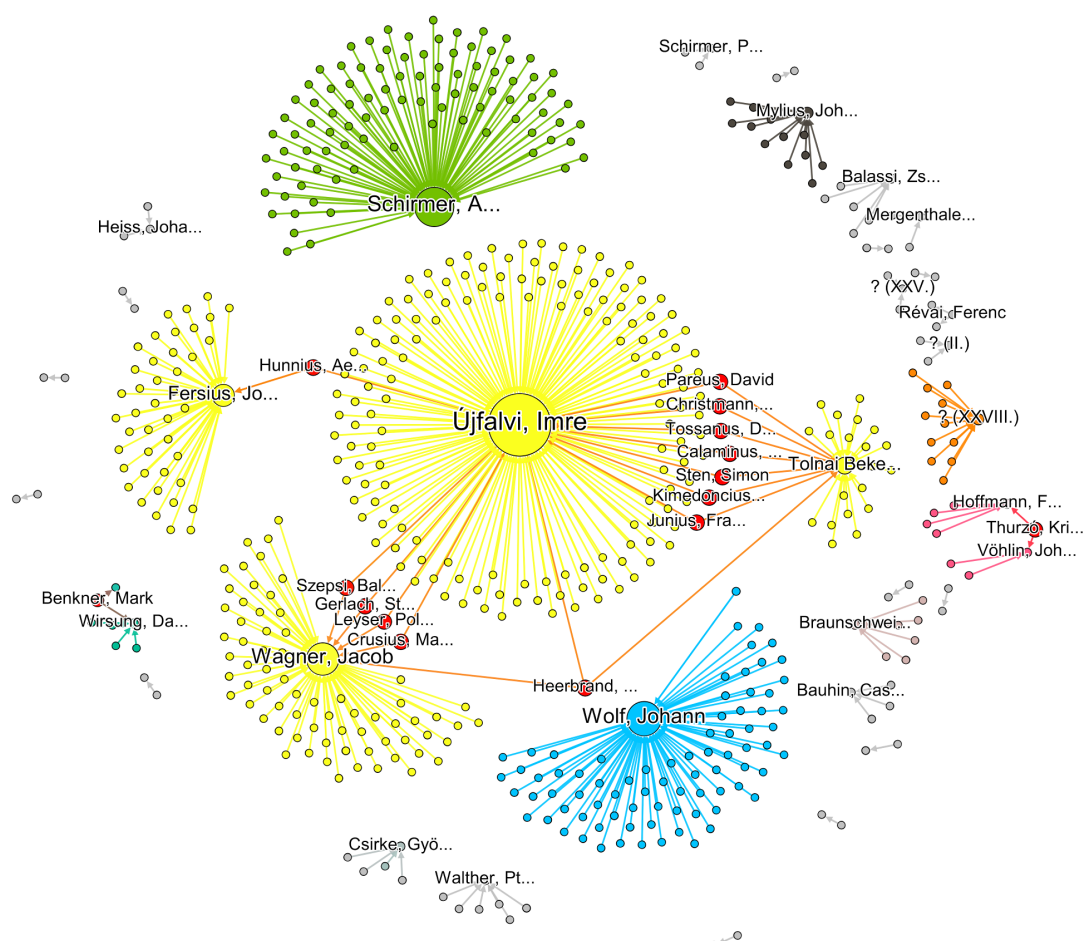
ID	Source	Target	Sign	Yeartype	Datum	Place
12117	Mandel, Christoph	Deyger, Johann Valentin	(Mh 1030)	16. század	1553. 06. 22.	Feuchtwangen*
12260	Lackhner, Leonhard	Hoffmann, Ferdinand	(Typ 515.47.456)	16. század	1555. -. -. ?	?
12264	Schwartz, Sebastian	Hoffmann, Ferdinand	(Typ 515.47.456)	16. század	1555. -. -. ?	?
12262	Thurzó, Elek	Hoffmann, Ferdinand	(Typ 515.47.456)	16. század	1555. -. -. ?	?
12261	Thurzó, Hieronymus	Hoffmann, Ferdinand	(Typ 515.47.456)	16. század	1555. -. -. ?	?
12263	Thurzó, Kristóf	Hoffmann, Ferdinand	(Typ 515.47.456)	16. század	1555. -. -. ?	?
12197	Thurzó, Kristóf	Vöhlín, Johann Jacob	(Egerton Ms. 1180)	16. század	1557. -. -. ?	?
12198	Thenn, Caspar	Vöhlín, Johann Jacob	(Egerton Ms. 1180)	16. század	1558. -. -. ?	?
12199	Welser, Anthon	Vöhlín, Johann Jacob	(Egerton Ms. 1180)	16. század	1558. -. -. ?	?
12106	Jordán, Tamás	Spon, Johann	(Egerton Ms. 1178)	16. század	1558. 01. 15.	Wittenberg
8118	Eber, Paul	Csirke, György	(Ant 7397)	16. század	1560. -. -. ?	Wittenberg*
8101	Major, Georg	Csirke, György	(Ant 7397)	16. század	1560. -. -. ?	Wittenberg*
8123	Moller, Heinrich	Csirke, György	(Ant 7397)	16. század	1560. -. -. ?	Wittenberg*
8034	Melanchthon, Philipp	Csirke, György	(Ant 7397)	16. század	1560. 01*. 01*. és 1560. 04*. 19*. között	Wittenberg*
11	Wacker, Jacob	Wirsung, David	(Stb. 134.)	16. század	1565 és 1583 között	Ambras
4445	Hofmann, Benedictus	? (XXVIII.)	(MS 675.)	16. század	1566 és 1568 között	?
4447	Mundilius, Leonhard	? (XXVIII.)	(MS 675.)	16. század	1566 és 1568 között	?
4444	Hermann, Simon	? (XXVIII.)	(MS 675.)	16. század	1566. 12. 27.	Wittenberg*

1. táblázat. A felhasznált adatok rendszere (részlet)

5.1. A bejegyzések hálózata 1500–1600 között

Az első vizsgált hálózat (1. ábra) a 16. századi bejegyzések rendszerét mutatja: a feldolgozott adatok a legelső 1553-as bejegyzéstől indulnak. A hálózat tartalmazza azon albumoknak az 1600 utáni bejegyzéseit is, amelyekben legalább egy darab 1600

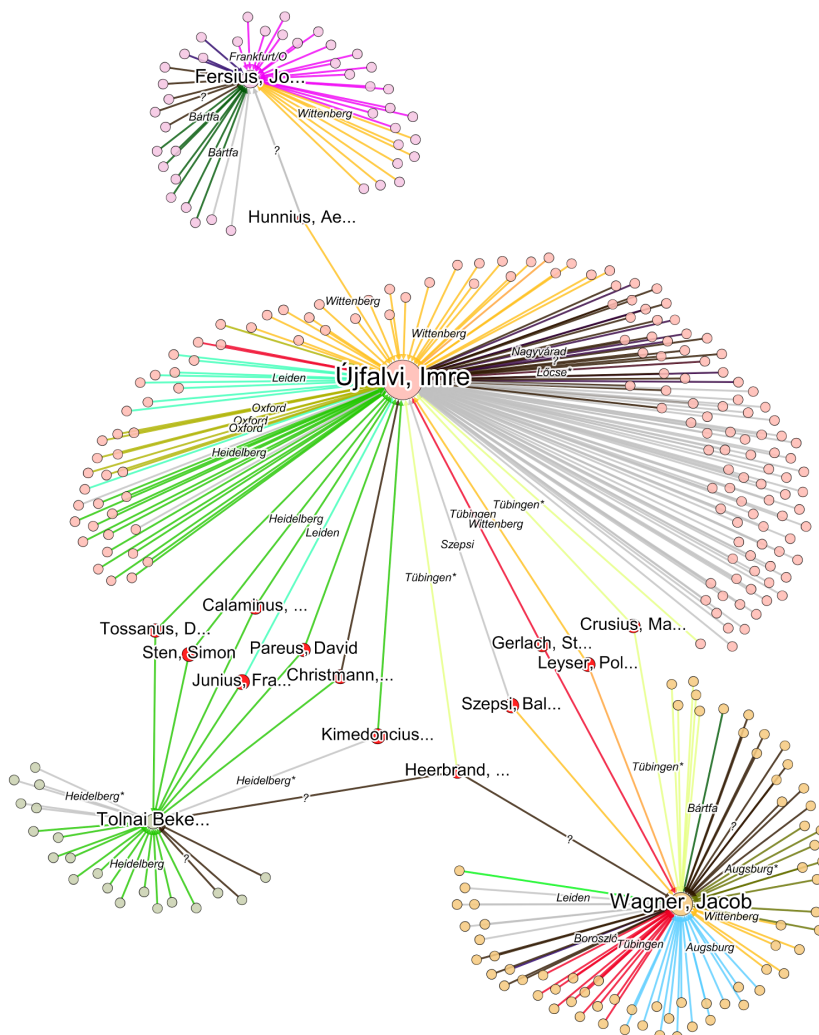
előtt keletkezett bejegyzés található. Ilyen módon az időbeli korlát nem töri meg a forrásokhoz kapcsolódó csoportokat.



1. ábra. Az 1500 és 1600 közötti bejegyzések hálózata

Az így kapott hálózatba 568 bejegyzés tartozik, 549 bejegyzőtől 36 címzettnek. A hálózat irányított, aszimmetrikus, tehát a kapcsolat (azaz bejegyzés) iránya mindig a bejegyzőtől mutat az inscriptio címzettje felé. A hálózat 585 pontból, személyből és 568 kapcsolatból áll. A hálózat nem összefüggő, 30 különálló komponensre bomlik. Az első ábrán szereplő hálózaton a különböző színek a különböző komponenseket jelölik, míg a második ábrán a bejegyzések helyszíne szerinti bontás látható. A komponensek nagyságának eloszlása, és a foksámeloszlás arra enged következtetni, hogy az albumbejegyzők közt is preferenciális kapcsolódás érvényesül. Meg kell jegyezni azonban azt is, hogy a preferenciális kapcsolódást idéző elrendeződés önmagából a forráscsoport jellegéből is adódik: ha valakinek van emlékkönyve, az egy tudatos gyűjtési folyamatban gyarapítja bejegyzői számát, az új bejegyzők tehát nem véletlen-

szerűen adódnak a hálózathoz. Az valószínűsíthető azonban, hogy aki sok bejegyzést gyűjtött magának, könnyebben, hatékonyabban tud új kapcsolatokat begyűjteni, mint azok, akik kevesebbet.



2. ábra. A hálózat a bejegyzések helye szerinti bontásban

Kérdéses, hogy az adatfelvétel módszere torzítja, erősíti-e ezt a preferenciális megjelenést: abból adódik-e a nagy fokszámú albumtulajdonosok jelenléte, hogy csak központi egyének albumai vannak feldolgozva, vagy hogy az ő albumuk feldolgozása részletesebb-e? A jelenség ennek a fordítottja is lehet: a nagy fokszámú elemeket épp azért tekinti központi figurának a szakirodalom, mert tőlük ismerünk nagyobb mennyiségű forrásanyagot. Sokkal kevesebb (magyar/hungarus) albumtulajdonost ismerünk, mint bejegyzőt, ez természetes és valós működésmódnak tűnik. Viszont a kisebb albummal rendelkezőket és az albummal nem rendelkező bejegyzőket is ismerjük, s míg egy a központiként ismert figurákat elemző kutatás a perifériára szorult szereplőket elkerülheti, addig a hálózati modellezésben meg tudnak jelenni

akkor is, ha nincs kapcsolódásuk az aktuális centrumhoz. Minden adatokkal és hiányos adatokkal dolgozó akár történeti, akár művelődéstörténeti kutatásnak szem előtt kell tartania, hogy az elemzés kizárólag a *jelenlegi tudásunk alapján* értelmezhető helyzetet mutatja.

Azok a csomópontok, személyek, akik nem rendelkeznek bemenő kapcsolatokkal, valószínűsíthetően nem rendelkeztek saját albummal, vagy albumuk elveszett. Tehát az, hogy a hálózat néhány nagy fokszerű komponensből, *ego-hálózat*ból áll, nem feltétlenül az adatgyűjtés jellegéből, hanem főként magából az adatok jellegéből következik. Akárcsak az, hogy az egyes bejegyzők között nincs (vagy nagyon ritka) a hungarika jellegű kapcsolat: hiszen ha két német bejegyző egymásnak írt az albumába, az a hungarika adatok között már nem szerepel.

A legtöbb bemenő kapcsolattal Újfalvi Imre rendelkezik, akinek kapcsolatrendszerét már évtizedek óta több kutatás is vizsgálta.⁴⁶ Míg 163 bejegyzést címeztek neki, addig arról még egyetlen forrást sem ismerünk, hogy ő bejegyzést dedikált volna ismerőseinek. A második legnagyobb komponens központi figurája Schirmer Anton, aki 96 bejegyzővel rendelkezik, de szintén nem írt bejegyzést. Johann Wolf 79, Jacob Wagner 74 bejegyzővel közel azonos magyar bejegyzői hálózatot működtet maga körül. Négy olyan csoport van, amelyben 10 és 50 közötti a bejegyzők száma, ez tekinthető a középmezőnynek; tizennégy olyan szereplő ismert, akinek csak egy bejegyzést címeztek. Közös jellemzőjük, hogy azok, akik bejegyzéseket fogadtak, nem írtak bejegyzéseket mások számára. Emellett úgy tűnik, hogy míg a szereplők törekedtek bejegyzéseket gyűjteni, addig kevésbé volt szándék minél több albumba írni: az egyes szereplők kimenő kapcsolatainak száma ugyanis jócskán alulmarad az átlagos bemenő fokszerűsokhoz képest.

Az eddig feltárt adatok alapján a hálózatban legaktívabb bejegyzőként Jacob Heerbrandt tűnik fel, ám ő is csupán 3 bejegyzést írt: egyet Wagnernek, egyet Újfalvinak és egyet Tolnai Beke Mihálynak. Rajta kívül csak Aegidius Hunnius írt három bejegyzést. Az ő esete különösen figyelemre méltó: ugyanis a három közül két bejegyzést ugyanannak a személynek, Johannes Fersiusnak címezte, néhány évnyi eltéréssel; míg a harmadik bejegyzése Újfalvihoz az egyetlen híd, amely összeköti a két alkotórészt. Jól látszik Hunnius összekötő szerepe a hálózatban. Emellett csak 15 olyan személy van, aki két bejegyzést is jegyzett. A legtöbb bemenő kapcsolattal az albumtulajdonosok rendelkeznek, ezt jelen esetben a presztízsértékük is tükrözi, tehát ők töltöttek be hatalmi pozíciót, azonban strukturálisan nem feltétlenül ők jelentik a hálózat legfontosabb pontjait, hanem azok a szereplők, akik a legtöbbféle szereplővel vannak kapcsolatban, s akik sok ágens közt jelentenek átmenő kapcsolatot. Újfalvi ilyen tekintetben is a legmagasabb értékekkel rendelkezik, de őt nem a többi magas fokszerű albumtulajdonos, hanem Hunnius és Heerbrand követi, illetve két egymással azonos helyzetben lévő bejegyzőkből álló összekötői csoport. Ezeket a hálózatban betöltött pozíciókat a második táblázat adatain is látni: Újfalvi magas fokszerűségéhez magas értékű közönségcentralitás is társul, de közelség centralitást tekintve már nem a második legnagyobb fokszerűséggel rendelkező Wagner következik, hanem Hunnius és Heerbrandt. Ők az információ áramlásának csomópontjai, ők vannak törésponti

⁴⁶ Keserű Bálint, „Újfalvi Imre és az európai »későhumanista ellenzék«,” *Acta Historiae Litterarum Hungaricarum* 9 (1969): 3–46.

pozícióban, akárcsak a velük hasonló helyzetben lévő (a külön csomópontokhoz kapcsolódó csoportokon belül egymással strukturálisan ekvivalens), piros színnel kiemelt bejegyzők. Ők strukturális hasonlóság alapján külön klasztereket alkotnak, ez szintén megfigyelhető a második táblázat bontásán.

A hálózat azonban nem túl sűrű: a lehetséges kapcsolatoknak csupán 0,17%-a valósul meg benne, s kevés az összekötő. Jellemzőbbek az egymástól elkülönülő komponensek. Nem jönnek létre úgynevezett *triádok*. (A *triád* 'háromszög, három személy közös kapcsolatrendszere, ahol mindenki mindenkivel kapcsolatban van, azaz hármas klikk.') Ha az 1500–1699-ig írott bejegyzéseket nézzük, ez változik: megjelenik egy óriáskomponens, amelyben a kapcsolatok az előbbinél jóval sűrűbbek, s már megtalálhatjuk a klikkesedés jelenségét is, tehát kialakulnak egymással teljes algráfot alkotó csoportot.

ID	fokszám	közelség-centralitás	közöttiség-centralitás	PageRank	triádok száma
Újfalvi, Imre	163	0,528	38287,000	0,123	0
Wagner, Jacob	74	0,360	18092,333	0,057	0
Fersius, Johannes	44	0,302	11571,000	0,034	0
Tolnai Beke, Mihály	28	0,324	5825,667	0,020	0
Heerbrand, Jacob	3	0,439	4893,875	0,002	0
Hunnius, Aegidius	3	0,385	10922,000	0,002	0
Gerlach, Stephan	2	0,414	2811,667	0,002	0
Crusius, Martin	2	0,414	2811,667	0,002	0
Szepsi, Balázs	2	0,414	2811,667	0,002	0
Leyser, Polykarp	2	0,414	2811,667	0,002	0
Calaminus, Petrus	2	0,364	521,208	0,002	0
Junius, Franciscus	2	0,364	521,208	0,002	0
Christmann, Jakob	2	0,364	521,208	0,002	0
Sten, Simon	2	0,364	521,208	0,002	0
Kimedoncius, Jakob	2	0,364	521,208	0,002	0
Tossanus, Daniel	2	0,364	521,208	0,002	0
Pareus, David	2	0,364	521,208	0,002	0
Clyne, David Johann	1	0,346	0,000	0,001	0
Gronbeccius, Isaac	1	0,346	0,000	0,001	0
Grynäus, Johann Jakob	1	0,346	0,000	0,001	0
Hosius, Johannes	1	0,346	0,000	0,001	0

2. táblázat. Az óriáskomponens töréspont-pozícióiban elhelyezkedő összekötők és a két (strukturálisan ekvivalens szereplőkből álló) összekötői csoport

5.2. Az 1500 és 1600 közötti bejegyzők hálózata szociomátrix alapján

A felhasznált adatok alapján elkészíthető a bejegyzők közötti kapcsolatok modellje. Az így elkészülő *szociomátrix* (3. ábra) a fentiekkel ellentétben nem azt mutatja, hogy ki kinek címzett *inscriptiót*, hanem azt, hogy mely bejegyzők találkoztak egy albumban. Tehát szimmetrikus kapcsolat áll fenn két bejegyző között, ha azok írtak bejegyzést



lálhatjuk: ez azt jelenti, hogy ő rendelkezik a legtöbb kapcsolattal, s ebből adódóan ő az a pont a hálózatban, aki a lehető legtöbb tagot könnyen elér. Míg Aegidius Hunnius kevesebb kapcsolattal rendelkezik nála, de sokkal több szereplő között jelenti ő az egyetlen átmenő utat: kiemelkedően magas a közöttségcentralitása. Sok út halad keresztül rajta, tehát strukturálisan hatalmi pozíciót birtokol, tipikusan ő az összekötő gyenge kapcsolat, aki az információáramlásban kiemelkedően fontos szerepet tölt be. Ezenfelül törésponton helyezkedik el: eltávolítása a hálózat két részre esését eredményezné. Utóbbi igaz Heerbandra és a többi kiemelt bejegyzőre is. Összekötő szerepet tölt be tehát mellettük két kiemelt csoport is, amelyek tagjai csoporton belül egymással strukturálisan ekvivalens szereplőkből állnak. (Tehát tagjaik a hálózatban való elhelyezkedésük szempontjából azonosak.) Az egyik ilyen csoport tagjai: Martin Crusius, Stephan Gerlach, Polycarp Leyser és Szepsi Balázs. Érdekes, hogy Szepsi Balázs, aki 1585-ben a wittenbergi egyetem diákja volt, s társszerzője egy 1586-os búcsúztatóversnek,⁴⁷ éppen olyan összekötő szerepet tudhat magának, mint amilyenre egyébként csak professzorok tudnak szert tenni a hálózatban.

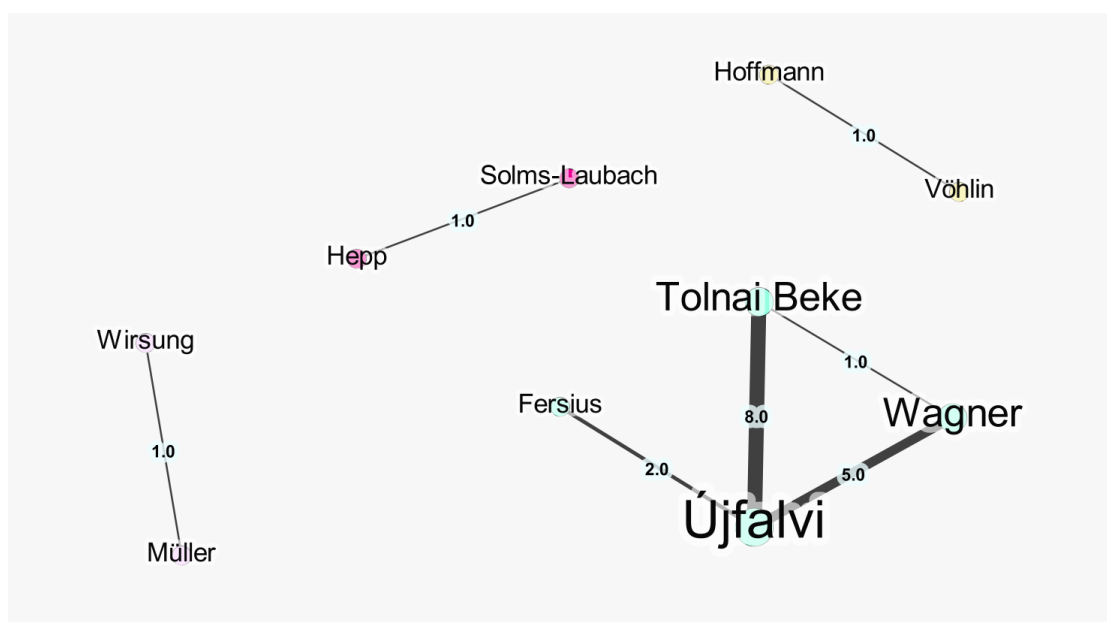
A másik csoportot Petrus Calaminus, Franciscus Junius, Jakob Christmann, Simon Sten, Jakob Kimedoncius, Daniel Tossanus és David Pareus alkotja. Mindkét csoportban az a figyelemre méltó, hogy a bejegyzők maguk közt is egy klikket, azaz teljes algráfot alkotnak, azaz minden tag kapcsolatban van minden másik taggal. Polycarp Leyser teológusként tanított, ahogy a tübingeni egyetem tanárai, Crusius és Gerlach is. Szepsi Balázs Wittenbergben volt diák 1585-től. Ők négyen Jacob Wagner albumán keresztül kapcsolódnak egymáshoz. A másik csoportban tulajdonképpen a heidelbergi egyetem tanári kara található: Calaminus, Junius, Tossanus, Pareus és Kimedoncius teológusként, Christmann orientalistaként és asztronómusként volt jelen Heidelbergben, Sten a filozófiai fakultás dékánja, majd az akadémia rektora lett. Kimutatható tehát, hogy összekötő szerepben elsősorban az egyetemeken tanító professzorok tudnak elhelyezkedni. Hunnius a legnagyobb összekötői értékkel rendelkező aktor a hálózatban, a centrumot pedig Heerbrand jelenti. Mindketten a wittenbergi egyetem professzoraként működtek: így lehettek kulcsfigurák az *album amicorum*ot használók számára.

5.3. Az 1500 és 1600 közötti albumtulajdonosok hálózata szociomátrix alapján

Az 1553 és 1600 közötti hálózatnak (1., 2. és 3. ábra) egy másik vetülete adható, ha azt nézzük, melyik albumtulajdonosnak voltak közös bejegyzői. A 4. ábrán látható, hogy a legtöbb címzettnek alapvetően nincsen kapcsolata egymással, elkülönülő pontokat alkotnak. Csupán négy olyan csoport van, amelynek a címzettjei között gyenge kapcsolatként egy vagy több bejegyző áll. Ezeknek a közösségeknek, alhálózatoknak a vizsgálata további kutatásokra lehet érdemes. Tolnai Beke Mihály, Újfalvi Imre, Jacob Wagner és Johann Fersius kapcsolati hálója emelkedik ki, hiszen ahogy az

⁴⁷ Frankl Vilmos, *A hazai és külföldi iskolázás a XVI. században* (Budapest: Atheneum, 1869), 314, hozzáférés: 2018.05.15, <http://mek.oszk.hu/12100/12146/12146.pdf>; Móré Tünde, „»Ars peregrinandi«: A 16. századi wittenbergi magyar peregrinusok neolatin búcsúztatóverseinek vizsgálata hazai és nemzetközi kontextusban” (Ph.D. disszertáció, Debreceni Egyetem, 2015), 165, https://dea.lib.unideb.hu/dea/bitstream/handle/2437/222188/MoreTteljesdissz_titkosított.pdf?sequence=1&isAllowed=y.pdf.

előbbi vetületen látható volt, ők jelentik az 1500–1600 közötti bejegyzések legnagyobb összekapcsolódó csoportosulását – egy heidelbergi és wittenbergi diákokat összekötő peregrinációs ismeretségi kört. Érdekes megfigyelni (2. ábra), hogy a Tolnait és Újfalvit összekötő hét, egymással strukturálisan ekvivalens szereplő mindegyik bejegyzése Heidelbergben datált – egy kivételével. A kapcsolati hálóban adódó egyezésekből talán feltételezhetjük azonban, hogy az eddig ismeretlen lokációval rendelkező bejegyzés is Heidelbergben keletkezhetett. Hasonló elv alapján felmerül, hogyha Tolnai albumának bejegyzései egy kivételével mind Heidelbergben keletkeztek, akkor a négy eddig ismeretlen helyzetű bejegyzésnek is lehet köze az egyetemvároshoz.

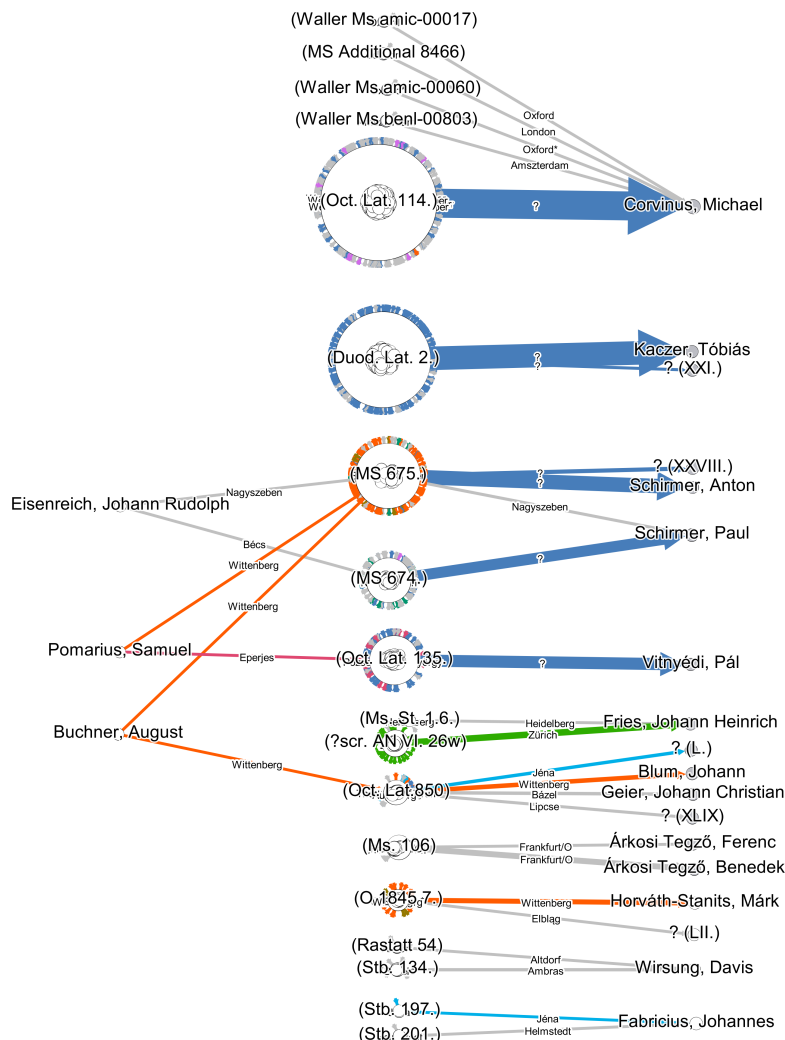


4. ábra. Az 1500 és 1600 közötti albumtulajdonosok hálózata szociomátrix alapján

5.4. Albumokhoz kötött háromosztatú (*tripartit*) hálózat

A forráscsoport lehetőséget ad egy *tripartit*, azaz háromosztatú (hárommódú) modellben való értelmezésre is (5. ábra). Ez egy olyan kapcsolathálót jelent, amelyben az elemek három különálló halmazra bonthatók: a különböző halmazok elemeinek egymással vannak kapcsolatai, adott halmazon belül az elemek azonban nem kapcsolódnak egymással. Az albumbejegyzések esetében ez azt jelenti, hogy felrajzolható egy olyan hálózat, amelyben a bejegyzők egy-egy albumhoz kapcsolódnak, az albumok pedig az albumtulajdonoshoz, címzetthez. Így egy a Bruno Latour-i értelemben vett cselekvő hálózatot kapunk, amelynek egyenértékű cselekvői nem csupán személyek, hanem tárgyak, esetünkben az albumok is lehetnek. A következőkben csak azokat az albumokat emeltem ki, amelyekben: a) több album kapcsolódik ugyanahhoz a címzetthez, tulajdonoshoz; b) ugyanahhoz az albumhoz több tulajdonos kapcsolódik. A háromosztatú hálózat közepén elhelyezkedő elemei az albumok, a kapcsolatok pedig elkülöníthetők a bejegyzés helye szerint. Az albumok és tulajdonosaik kapcsolatának helyszíne az album összes bejegyzésének legtöbbször előforduló helyszínmegjelölését

mutatja. A vizualizációban a forrásokat jelölő halmazok magukban tartalmazzák a bejegyzőiket – csupán azokban az esetekben emeltem ki a bejegyzőket, amikor azok összekötő szerepet töltenek be.

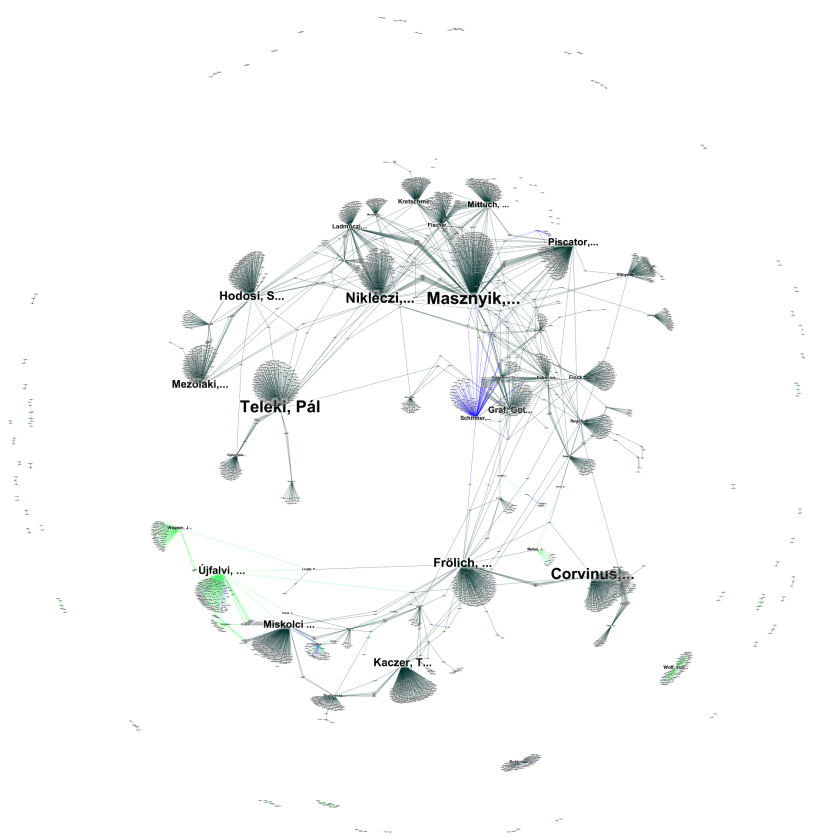


5. ábra. Az albumok, a bejegyzők és a tulajdonosok háromosztatú hálózata

A jelenleg elérhető adatok alapján⁴⁸ a különböző albumok általában különböző helyeken keletkezett bejegyzéseket tartalmaznak: Michael Corvinusnak például saját, 210 bejegyzést tartalmazó albuma mellett még 4 másik albumban tűnik fel a neve címzettként egy-egy bejegyzésben. Ezek közül egy oxfordi, egy feltételezhetően oxfordi, egy londoni és egy amszterdami helymegjelöléssel készült. Johann Heinrich Friesnél⁴⁹ szintén arra látunk példát, hogy sok bejegyzést tartalmazó albumába csak Zürichben

⁴⁸ Az IAA-ból átvett adatokon túl az adatbázisban található egyéb információk az elemzéskor nem voltak elérhetőek.

⁴⁹ Friesről bővebben: Hegyi Ádám, „A gályarab lelkészek bejegyzései Johann Heinrich Fries (1639–1718) albumában: magyarországi diákok Bázélben és Zürichben 1677 és 1720 között,” *Egyháztörténeti*



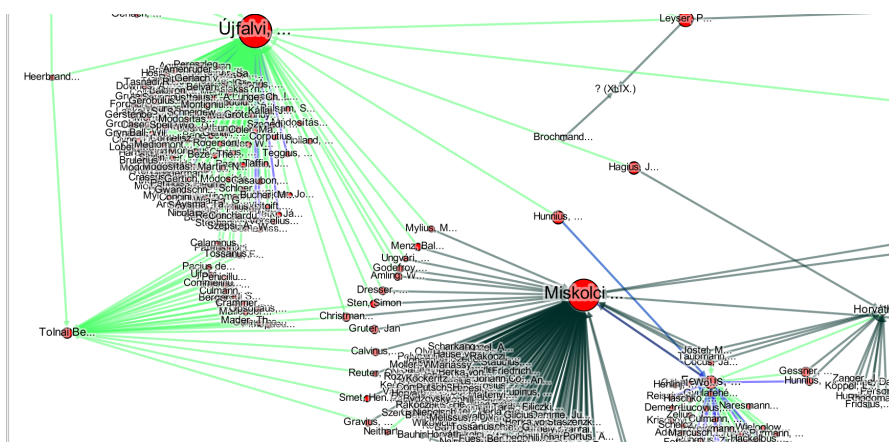
6. ábra. A bejegyzések hálózata 1500 és 1700 között⁵⁰

⁵⁰ A nagyfelbontású változat elérhető a tanulmány mellékleteként.

ezen a komponensen végezni, megnézni, kik az összekötők, miként oszlik fel a hálózat a bejegyzések helye szerint, mely klikkek működnek.

Az óriáskomponens kapcsolati sűrűsége valamivel nagyobb, mint a teljes hálónak, az 1500–1600 közti közösség sűrűségétől pedig messze elmarad. Pedig – szubjektív megközelítéssel – ránézésre egy szövedékesebb hálózat lehetne. Azonban míg az elemek száma majdnem hatszorosára növekedett, addig a tendencia, hogy a bejegyzők maximálisan csak egy vagy két bejegyzést írnak, megmaradt. Az óriáskomponens jelenségeire koncentrálva az is megállapítható azonban, hogy bár a sűrűség kicsi, de az összekapcsolódás minősége jelentősen változott. Míg a megelőző időszakban egyetlen klikket sem találtunk, jelen esetben a *triádok*, háromszög-kapcsolatrendszerek száma 132-re nőtt.

A megnövekedett hálózatban már jól látszik, hogy a legnagyobb fokszerű pontok nem feltétlenül jelentik a hálózatot összetartó pontokat. Hiába tűnik a legfontosabb szereplőnek Masznyik Tóbiás, aki 229 bejegyzéssel a legnagyobb album tulajdonosa lehetett, mégis a közönségcentralitást nézve csak a negyedik legmagasabb értékkel rendelkezik. (Pontos adatok: 3. és 4. táblázat.) Közvetítő szerep tekintetében sokkal erősebb pozícióban van David Fröhlich és Nikléczi Boldizsár. A hely szerinti bontás (8. ábra) segítségével pedig az is megállapítható, hogy a hálózat fő vonulatát a wittenbergi keltezésű bejegyzések alkotják, a korábbi hálózatot a nagyobb, 1700-ig tartó hálózatba bekapcsoló bejegyzések is wittenbergi, illetve heidelbergi kapcsolatokat jelentenek. Ki kell emelni azt is, hogy presztízs alapján, amely nemcsak a kapcsolatok számát, hanem értékét és számításba veszi, Teleki és Corvinus áll a bejegyzési hálózat centrumában.



7. ábra. Az óriáskomponens és az összekötők (nagyításban)

A 9. ábrán megfigyelhető, hogyan néz ki az albumtulajdonosok hálózata, amelyben a kapcsolatot a közös bejegyző jelenti. Jól látszik, hogy különböző ismerettségű körök különíthetők el, amelyeket több esetben is néhány összetartó, magas közöttségcentralitással rendelkező szereplő (például Jarkovitzky Dániel, Hoffmann Ferdinánd vagy éppen Miskolczi Csulyak István) tart össze. Az így kimutatható ismerettségű körök vizsgálata izgalmas megfigyeléseket eredményezhet, jelen tanulmány keretei között erre azonban nincs lehetőségem.

A 6. és 7. ábrákon a színek szerinti lebontás a hálózat növekedését mutatja: a zöld élek az 1500–1600 között már meglévő alhálózatot jelzik, a világoskékek azokat a

kapcsolatokat, ahol a 16–17. század fordulóján, vagy nem meghatározott időpontban készültek a bejegyzések, a sötétkéék vonalak pedig a 17. századi bejegyzéseket jelölik. Érdekes megfigyelni, hogy Johann Wolf albumának bejegyzői különálló csoportban maradtak. A hálózat növekedése Újfalvi Imre, Johannes Fersius, Tolnai Beke Mihály, Anton Schirmer és Johannes Mylius ismeretségi körének köszönhető.

ID	fokszám	közelség-centralitás	közöttiség-centralitás	presztízs	triádok száma
Maszyik, Tóbiás	229	0,241	950054,8	0,026	5
Teleki, Pál	224	0,217	890272,0	0,029	0
Corvinus, Michael	215	0,190	698539,4	0,029	0
Nikléczi, Boldizsár	206	0,244	1191356,0	0,023	42
Frölich, David	182	0,232	2054732,2	0,023	0
Hodosi, Sámuel	177	0,202	673851,4	0,022	3
Kaczer, Tóbiás	167	0,190	558851,8	0,022	0
Újfalvi, Imre	163	0,173	549965,0	0,021	0
Mezőlaki, János	162	0,187	509687,4	0,021	0
Miskolci Csulyak, István	157	0,169	533668,0	0,020	1

3. táblázat. Az óriáskomponens legnagyobb fokszámú szereplői

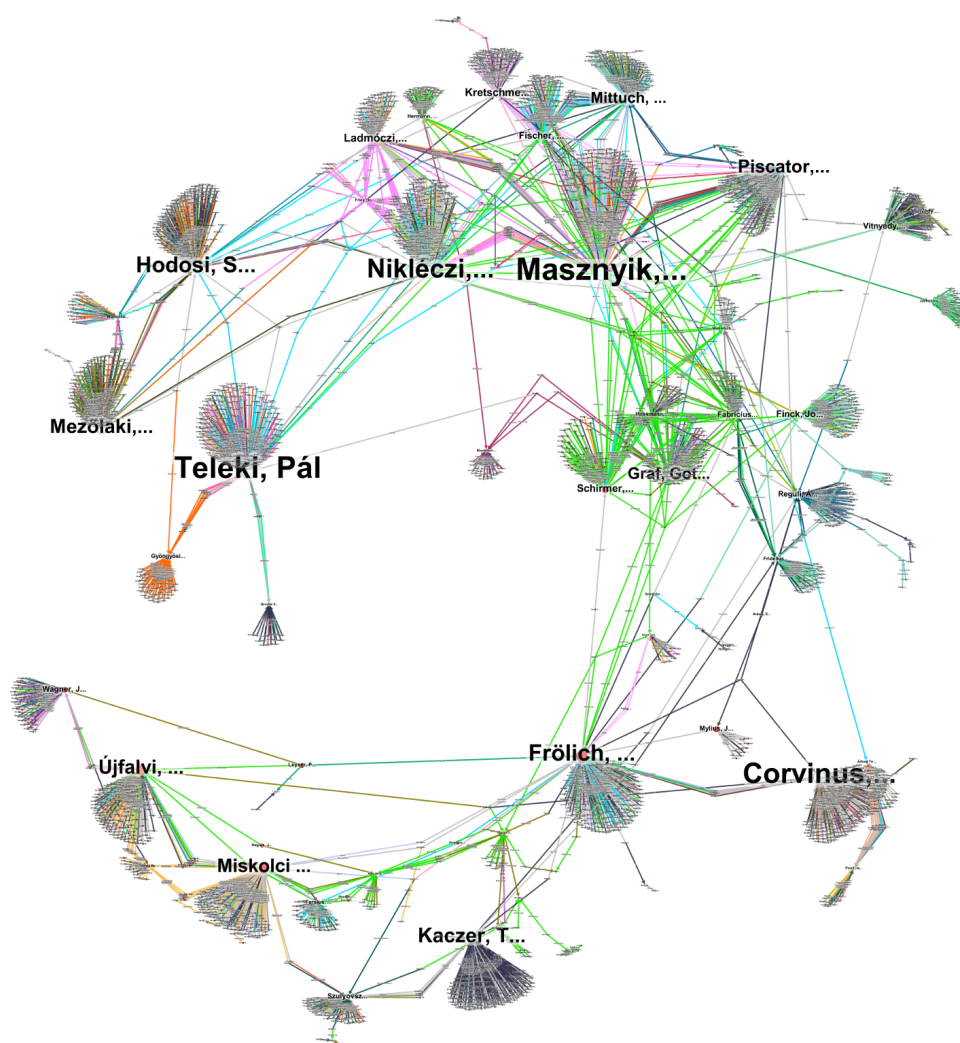
ID	fokszám	közelség-centralitás	közöttiség-centralitás	presztízs	triádok száma
Frölich, David	182	0,232	2054732,2	0,023	0
Nikléczi, Boldizsár	206	0,244	1191356,0	0,023	42
Graf, Gottfried	139	0,255	1176223,2	0,015	15
Maszyik, Tóbiás	229	0,241	950054,8	0,026	5
Teleki, Pál	224	0,217	890272,0	0,029	0
Reguli, Ábrahám	71	0,240	834865,6	0,009	0
Scharff, Johann	6	0,247	713718,1	0,001	2
Corvinus, Michael	215	0,190	698539,4	0,029	0
Piscator, Nicolaus	154	0,226	675259,0	0,018	0
Hodosi, Sámuel	177	0,202	673851,4	0,022	3

4. táblázat. Az óriáskomponens legnagyobb közöttiségcentralitású, töréspont-pozícióiban elhelyezkedő szereplői

6. Összegzés

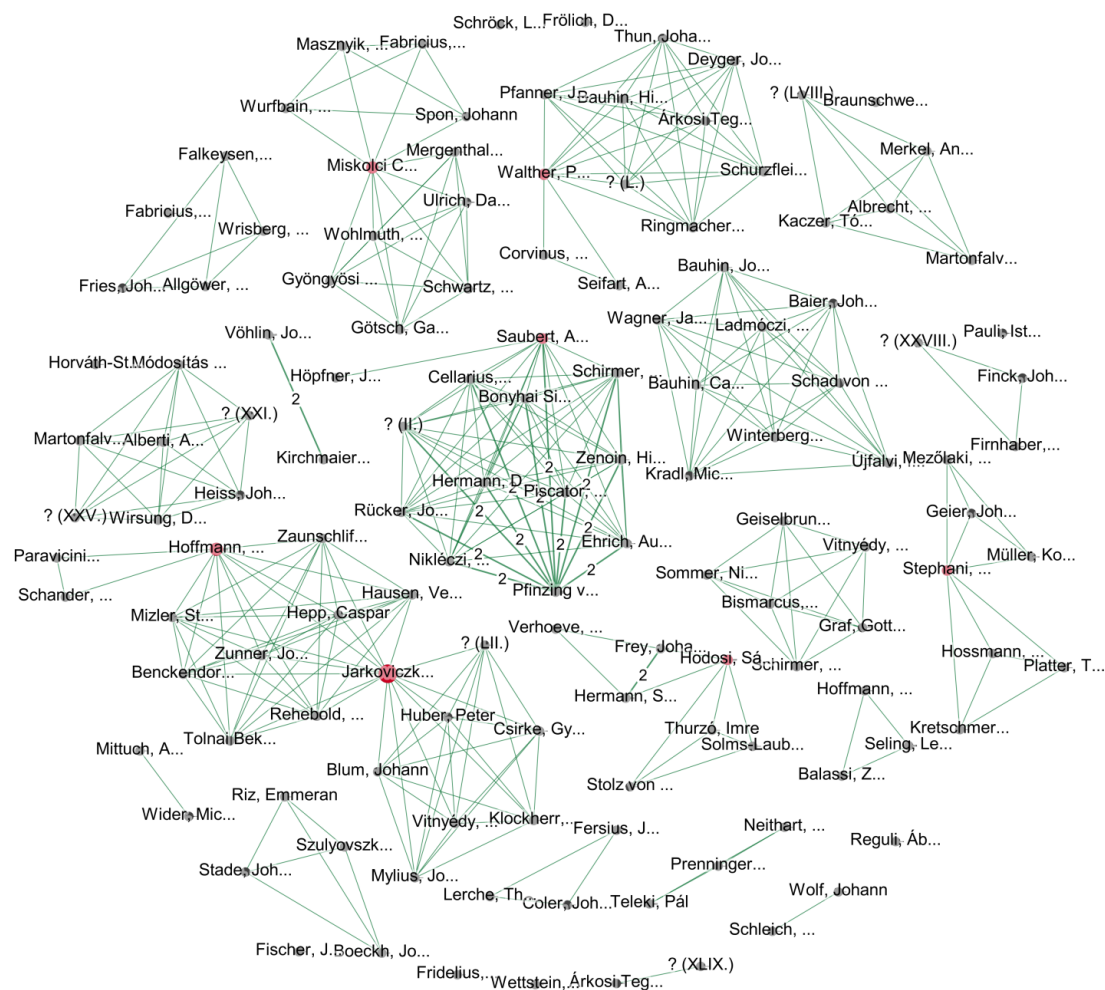
A tanulmány célja elsősorban az emlékkönyvbejegyzések (bejegyzők és albumtulajdonosok) hálózatként való értelmezése volt. Az IAA-ban feldolgozott adatok egy hálózati modellként interpretálhatók, rajzolhatók fel és értelmezhetők, amelyben megfigyelhetők és vizsgálhatók a szociális hálózatokra is jellemző tulajdonságok. A módszertani kísérlet ilyen szempontból tehát működőképesnek látszik: a hiányos adatok ellenére

is az ismert tényezők mintegy véletlen mintavételként következtetni engednek a hálózat működési tendenciáinak nyomaira; a jelenleg ismert és adatolt bejegyzők, tulajdonosok és bejegyzéseik pedig indikátorként egy a továbbiakban bővíthető struktúra alapjaiként szolgálhatnak. A hálózattudomány és a társadalmi hálózatelemzés matematikai és informatikai módszertani kerete és eszközei adaptálhatóak az album-bejegyzések művelődéstörténeti korpuszára, bár azt nem szabad szem előtt tévesztetni, hogy az adatok hiányossága miatt ez nem reprezentatív mintavételt jelent, s ezért a következtetések mindig a jelenleg ismert vetületről adnak információt. Az elemzés másik kérdésére, miszerint feltárható-e a bejegyzők közötti, bejegyzések általi kapcsolódások mintázata, s hogy a megközelítésmód hozzájárulhat-e eddig rejtőző jelenségek megvilágításához, kimutatásához – a választ a következőkben összegzem.



8. ábra. A hálózat óriáskomponense a bejegyzések helye szerinti bontásban

Az IAA adatbázisból összesen 4080 bejegyzésre vonatkozó adatot dolgoztam fel, az elemzések az 1553-as, legelső bejegyzés és az 1692–1740-as bejegyzéscsoport közötti



9. ábra. A tulajdonosok hálózata közös bejegyzőik alapján

intervallum adatait tartalmazzák. Több időkeretben és bontásban is vizsgáltam az adatokat: az első hálózat (1. ábra) az első 50 év adatait tartalmazza, a nagyobb időkeret pedig már azt mutatja, hogy egy évszázaddal később milyen arányú a kapcsolatháló növekedése az ismert bejegyzések által (6. ábra). Az 1553–1600 közötti hálózat időkerete ad releváns megközelítést a szociális háló konkrétabb vizsgálatához, ezért az akkori bejegyzőknek és tulajdonosoknak több hálózatát is megvizsgáltam: a bejegyzések által összekötött gráfon túl a szereplők *szociomátrixa* is kidolgozható. Ez a bontás tipikusan a gyenge kapcsolatok detektálását segíti, és megmutathatja az adatok listázásával, statisztikai vizsgálatával nem elérhető, megbújó információkat. Jelen esetben azt a gyenge kapcsolatot mutatja meg, amelyben két bejegyző kapcsolódik egymáshoz, ha mindketten szerepelnek adott albumban (4. ábra); illetve azt a kapcsolatot, amelyben két albumtulajdonosnak közös bejegyzője van (3. ábra). Így például kimutatható több olyan közösség, alhálózat is, amelyek vizsgálata további kutatásokra lehet érdemes. Az

egyik leginkább összekapcsolódó komponens Tolnai Beke Mihály, Újfalvi Imre, Jacob Wagner és Johann Fersius kapcsolati hálójá: a hálózat ismert szereplői közül ők alkotják az 1500–1600 közötti bejegyzések legnagyobb összekapcsolódó csoportosulását – egy heidelbergi és wittenbergi diákokat összekötő peregrinációs ismeretségi kört.

Érdekes lehet a bejegyzések hálózatban levő pozícióját és keletkezési helyük kapcsolatát is megvizsgálni (2. ábra): például a Tolnait és Újfalvit összekötő hét egymással strukturálisan ekvivalens szereplő mindegyik bejegyzése Heidelbergben datált – egy kivételével. A kapcsolati hálóban adódó egyezésekből azonban talán feltételezhetjük, hogy az eddig ismeretlen lokációval rendelkező bejegyzés is az egyetemvároshoz köthető. További vizsgálatokra adna lehetőséget, ha átvehetők lennének és a bejegyzésekhez hozzákapcsolhatóvá válnának magára az inscriptio szövegére vonatkozó információk (az inscriptio nyelve, kitől való idézet, stb.), és így elemezhető lenne az idézetek közlekedése is a jelen vizsgálatban felrajzolt és kimutatott hálózatban.

Az adatok hiányából, és az adatok szerkezetéből fakadó torzítás ellenére van néhány olyan jelenség, amely kimutatható az első 50 év hálózatában, majd a hálózat növekedésekor is; így talán feltételezhető, hogy egy létező mintázat nyomaira utal. Ilyen mintázatként figyelhető meg, hogy bár a legtöbb bemenő kapcsolattal az albumtulajdonosok rendelkeznek, azonban strukturálisan nem feltétlenül ők a hálózat legfontosabb pontjai. Azok a bejegyzők lesznek az információ áramlásának csomópontjai, akik a legtöbb szereplővel képeznek kapcsolatot: ők vannak törésponti pozícióban. Így például az 1553 és 1600 közötti bejegyzések között jelenleg Újfalvit ismerhetjük a legnagyobb összekötő értékekkel rendelkezőként, ám őt nem a többi magas fokszerű albumtulajdonos, hanem Hunnius és Heerbrand követi, illetve két egymással azonos helyzetben lévő bejegyzőkből álló összekötői csoport. A hálózati modellben tehát kimutatható, hogy összekötő szerepben elsősorban az egyetemeken tanító professzorok tudnak elhelyezkedni, így kulcsfigurák lehettek az *album amicorum*ot használók számára.

A hálózat növekedése során érdemes arra figyelni, hogy miként változik a hálózat struktúrája, az összekapcsoltság ugyan számszerűen nem növekszik – de a sok apró, periférikus bejegyző-címzett pároson kívül kialakul egy centrális óriáskomponens, amelyben már nemcsak egymástól elkülönülő énhálózatok laza szövedékét látjuk, hanem elkezdődik egyfajta klikkesedés. Míg az első ötven év ismert adatai között egyetlen három vagy több fős egymással teljesen összekapcsolt csoportot, klikket sem találunk, addig ha az 1500–1699-ig írott bejegyzéseket nézzük, ez változik: a klikkesedés megjelenése fontos tényező az összekapcsoltság minőségének szempontjából. Mindezen felül a 9. ábra megmutatja, hogy milyen ismeretségi körökre bomlik az *album amicorum*ot használók hálózata a 16–17. században.

A hálózati modell lehetőséget ad arra, hogy új kontextusban láthassuk az *album amicorum*ok eddig ismert bejegyzéseit, rátaláljunk olyan kiemelt értékű szereplőkre, elkülöníthető csoportokra, amelyek egy nagyobb, lineáris elrendezésű adathalmazban megbújhatnak. A hálózati modellben való megközelítés rámutat arra, hogy a fent említett szereplőket és összekapcsolódó csoportokat érdemes közelebbről is szemügyre venni és további vizsgálatokat folytatva összekapcsolódásuk tartalmi elemeit megvizsgálni.

Network Analysis of Album Amicorums in the 16th–17th Centuries: The Network of the Hugarica Inscriptions in the Autograph Books Between 1500 and 1700 Based on the *Inscriptiones Alborum Amicorum* Database

The inscriptions, personal greetings of the *album amicorums* in the 16th–17th centuries can be seen as remarkable resources for several research projects: philological, historical or biographical data can be identified; as well as the flow of knowledge can be detected. These, autograph books are the early examples of networking, and the set of these citations and signatures, included in these books, offers the possibility to interpret them as a network-model. In my study I would like to analyse the inscriptions of the autograph books between 1500 and 1700 as such a network that includes the owners and the inscriptors of the albums as nodes and the inscriptions as edges, contacts. This analysis would like to experiment with or test the method of network science and network analysis on an early modern corpus. In this study I will answer the questions how these interdisciplinary methods of network analysis could be adaptable to a culture-historical corpus and whether this method could help to reveal new or hidden philological or biographical nuances.

Keywords:

album amicorum, autograph books, intellectual networks, early modern, network analysis

Matthew L. Jockers

Washington State University, Pullmann, WA, College of Arts & Sciences

mjockers@unl.edu

Metaadat*

A *Macroanalysis: Digital Methods and Literary History* című könyv 5. fejezete azt kísérli meg bebizonyítani, hogy a bibliográfiai és demografikus metaadatok számítógépes módszerekkel történő elemzése lehetővé teszi irodalomtörténeti narratívák újraértékelését vagy átírását. A fejezet bemutatja, hogy irodalomtörténeti korszakok és trendek gondolatébresztő perspektívákba helyezhetők a könyvszintű metaadatok – könyvcímek, szerző származása, publikáció dátuma, kitalált helyszín és időpont stb. – makroelemzésével. Az ír-amerikai irodalomhoz kötődő metaadat-adatbázis hasznosításával a szerző újraértékeli az eddig elfogadott ír-amerikai irodalomtörténeti narratívát és egy alternatív perspektívába állítja Charles Fanning elméletét az ír-amerikai írók „elveszett generációjáról.” A metaadatok kontextust teremtenek Fanning irodalomtörténeti olvasatához, és azt sugallják, hogy az ír-amerikai irodalom történetére vonatkozó tudományos feltételezések a szerzők egy homogén csoportjának kisszámú művének elemzésén alapulnak. Absztraktabban fogalmazva a fejezet mellett érvel, hogy a hagyományos irodalomtudósok tévesen gondolják, hogy a nagy merítésen alapuló elemzések helyettesíteni szeretnék a szoros olvasást. Éppen ellenkezőleg, a metaadatok makroelemzése kizárólag a szükséges kontextust teremti meg a szoros olvasáshoz, és új kérdések felvetéséhez, az irodalomtörténet új perspektíváinak megalkotásához járul hozzá.

Kulcsszavak:

makroanalízis, metaadat, távoli olvasás, szoros olvasás, ír-amerikai irodalom



Hogy megválaszolhasd azokat a kérdéseket, jó metaadatokra van szükséged.

Geoff Nunberg

Arra, hogy a makroanalitikus megközelítés irodalomtörténeti ismereteinkhez miként járul hozzá újfajta tudással, e fejezet nyújtja az első példát. Egyúttal e fejezettel kezdődik az a szélesebb körű hatásvizsgálat is, amely vezérfonálként húzódik végig a

* *Macroanalysis: Digital Methods and Literary History*. (Urbana: University of Illinois, 2013), 35–62 (Chapter 5. Metadata). Copyright 2013 by the Board of Trustees of the University of Illinois. Used with permission of the University of Illinois Press. A szöveg fordítását az EFOP-3.6.1-16-2016-00008 azonosítójú, EU társfinanszírozású projekt támogatta.

könyvön. Az itt bemutatott bizonyíték elsősorban kvantitatív; egy nagy méretű irodalmi bibliográfiából gyűjtöttük ki egy erre a célra készült számítógépes eszközzel. A fejezet bizonyos mértékben az irodalomtörténet legalacsonyabban termő gyümölcsének leszüreteléséről szól. A tömeges digitalizáció előtt néhány évtizeddel a könyvtárak online elérhető elektronikus katalógusok formájában már digitalizálták gyűjteményeik egyik fontos összetevőjét. Ezek a kereshető bibliográfiák metaadatok formájában sok információt tartalmaznak. Gondoljunk például arra, hogy a Kongresszusi Könyvtár raktári számai és tárgyszófejlécei mit jelenítenek meg? Az irodalmár kutatók tudják, hogy a „P” sorozat különösen releváns munkájuk során, és hogy a „PR”-rel vagy „PS”-sel jelzett munkák még inkább relevánsak – ez az angol nyelv és irodalom. Mindez egy bőséges, még ha általános formája is a feldolgozható, kibányászható irodalmi adatnak. A tárgyszófejlécek még gazdagabb források. A fejléceket humán kódolók készítették, akik annak érdekében, hogy meghatározzák, a mű fikciós vagy nem fikciós, hogy az alkotás népköltészeti vagy angol reneszánsz, az amerikai irodalom esetében pedig, hogy a szöveg az északi, a déli, a középső vagy a nyugati régióból származik: időt szakítottak az általuk katalogizált szöveg ellenőrzésére.

A katalógusok ilyen típusú metaadatai teljes mértékben kiaknázatlanok voltak az irodalomtörténeti vizsgálatok során. Még az irodalmi bibliográfiák készítői is inkább csak arra fókuszáltak, hogy minél kimerítőbb bibliográfiákat állítsanak össze, mintsem arra, hogy a benne foglalt adatok felhasználhatók volnának irodalomértésünk új tudással való bővítésére. A bibliográfiai metaadat teljes szöveg hiányában is hasznos információkkal szolgál az irodalmi trendekről. 2003-ban Franco Morettivel két bibliográfiai adatkészletet használva vizsgálatsorozatba kezdtünk. Moretti adatkészlete 19. századi regények bibliográfiája volt: címek, szerzők és kiadási dátumok; nem részletes metaadatok, de számos tétel, körülbelül 7000 megidézett szöveg. Moretti munkája végül a 19. századi regénycímekről szóló tanulmányhoz vezetett, ami a *Critical Inquiry*-ben jelent meg.¹ Az én adatkészletem sokkal kisebb volt, egy ír-amerikai szerzők nagyjából 800 munkáját tartalmazó gyűjtemény.² Az ír-amerikai bibliográfia mindazonáltal alaposan gondozott, amelyet metaadatokkal bővítettünk, ezek a mű földrajzi elhelyezését jelölték, valamint a szerző nemét, születési helyét, korát és lakhelyét. A tételhez földrajzi koordinátákat – hosszúsági és szélességi fokok – is hozzárendeltek, amelyek azt jelölték, az egyes szerzők honnan származtak, az egyes műveket hol készítették.

Az ír-amerikai adatbázis az Egyesült Államok nyugati részének ír-amerikai irodalmával foglalkozó disszertációm háttérkutatásaként indult.³ 2001-ben az elsődleges források eredeti bibliográfiáját egy kereshető relációs adatbázisba vittem be, ami tá-

¹ F. Moretti, „Style, Inc.: Reflections on Seven Thousand Titles (British Novels, 1740–1850),” *Critical Inquiry* 36, 1. sz. (2009): 134–158.

² Moretti projektjének eredményeit először a 2007-es Champaign-Urbana-i digitális bölcsészettudományi konferencián ismertette, majd a *Critical Inquiry*-ben publikálta. Az ír-amerikai adatkészleten végzett elemzéseim eredményeit két különböző formában mutattam be: először 2007-ben a University of St. Louis meghívott előadójaként „Metadata Mining the Irish-American Literary Corpus” címmel, majd ismét 2007-ben a Modern Language Association összejövetelén, cikkformában *Beyond Boston: Georeferencing Irish-American Literature* címen.

³ M. Jockers, „In Search of Tir-Na-Nog: Irish and Irish-American Literature in the West” (Ph.D. diss., Southern Illinois University, 1997).

mogatja a gyors és könnyű keresést, sorba rendezést. Az adatbázisba kerülő művek kiválasztásának kritériumait – néhány kisebb változtatással – Charles Fanning nagy jelentőségű, az ír-amerikai irodalom történetéről szóló, *The Irish Voice in America: 250 Years of Irish-American Fiction* című munkájából kölcsönöztem.⁴ Ahhoz, hogy az adatbázisba kerülhessen, egy írónak rendelkeznie kellett néhány igazolhatóan ír összel, munkájának arról kellett szólnia, foglalkoznia kellett azzal, milyen írnek lenni Amerikában. E második kritérium miatt néhány nyilvánvalóan ír szerző, mint például F. Scott Fitzgerald és John O'Hara, nem szerepel a gyűjteményben. Amint Fanning és mások érvelnek, ők ketten ír gyökereiktől távolságot akartak tartani, ezért kerülték, hogy etnikai kérdésekről írjanak. Ennélfogva az adatbázis nem pusztán az ír gyökerekkel rendelkező írókra összpontosít, hanem azokra az ír gyökerekkel rendelkező írókra, akik prózájukban az ír identitás feltárását tették témává.

Annak meghatározása, hogy egy művet miként kell felvenni az adatbázisba, időnként szubjektív döntések eredménye volt. A meghozott döntések némelyikét meg lehet kérdőjelezni, és talán meg is kell. Tökéletes példa erre Kathleen Norris kaliforniaiként való rögzítése. Norris a San Franciscó-i ír közösségben nőtt fel, és itt kezdte írói pályáját is. Azonban miután megházasodott, New Yorkba költözött, ahol folytatta az írást, fikciós művei pedig hol Kaliforniában, hol New Yorkban játszódnak. Az adatbázis következetesen kaliforniaiként és nem New York-iként osztályozza. Norrist erőteljesen meghatározta San Franciscó-i ír-amerikai öröksége, ahogy önéletrajzában állítja, a San Franciscó-i „írókrácia” tagja volt.⁵ Annak ellenére, hogy jó ideig a keleti parton élt, soha nem veszítette el kaliforniai ír identitását. Végül visszatért Kaliforniába, és Palo Alto-i otthonában hunyt el.

Ez a tíz éven át épített adatbázis az ír-amerikai próza 758 művének, 250 évének bibliográfiai adatait tartalmazza.⁶ A nagy időtáv ellenére mégsem hatalmas korpusz. Mindazonáltal olyan korpusz, amely a kiválasztás kritériumait tekintve a teljességre törekszik. Az adatbázis minden tétele egy teljes bibliográfiai leírást tartalmaz, egy rövid kivonatot és kiegészítő metaadatokat, amelyek a könyv készültére vonatkoznak: földrajzi koordináták és olyan információk mint az állam, a régió, valamint olyan szubjektívebb módon kinyert információk, mint az, hogy a szöveg elsősorban városi vagy vidéki környezetben játszódik-e? Az adatbázistételek továbbá a könyv szerzőjéről is tartalmaznak információkat: nemüket és néhány esetben egy rövid életrajzi kivonatot is.⁷

⁴ C. Fanning, *The Irish Voice in America: 250 Years of Irish-American Fiction*, 2nd ed. (Lexington: University Press of Kentucky, 2000).

⁵ K. T. Norris, *Family Gathering* (Garden City, N.Y.: Doubleday, 1959).

⁶ Az adatbázis a széppróza mellett más típusú műveket is tartalmaz, de az elemzés során kizártam ezeket. Doktori disszertációm részeként kezdtem gyűjteni ezeket az adatokat, és egészen 2005-ig folytattam a bővítést. Ez idő alatt a Stanfordon tartott ír-amerikai kurzusaim diákjai segítettek, valamint olyan hallgatók, akiket a Stanford Humanities Labtól elnyert, „Irish-American West” című projektből alkalmaztam. 2004-ben az ír konzullal, Dónal Denhammal együtt társalapítója voltam a Western Institute of Irish Studies-nak, és nekik adtam az adatbázist, hogy tegyék nyilvánosan elérhetővé. 2009-ben lemondtam az intézet vezetéséről, és néhány évvel később a projekt lezárult, az adatbázis többé már nem elérhető.

⁷ Az adatoknak elemzés céljára történő emberi „kódolás” folyamata érdekes eredményekhez vezetett Gotschallnál. J. Gotschall, *Literature, Science, and a New Humanities* (New York: Palgrave Macmillan, 2008).

Az adatbázis hiányában az amerikai próza egyik egyedi etnikai alműfajáról, az ír-amerikai irodalomról szóló tudásunk most már csak egy-egy kánonba került szerzőről szóló kritikai tanulmánytól és Fanning e tárgyú munkájától függ.⁸ Fanning a *The Irish Voice in America*-ban a kánon történetét és fejlődését „generációs” megközelítést alkalmazva tárja fel. Az ír történelem vízvonalozó eseményével indít, az 1840-es évek ír éhínségével, ezt követően mutatja be az éhség előtti, alatti és utáni írók generációit. A 20. század fordulójához érve az amerikai és a világtörténelem kulcsfontosságú eseményeinek kontextusára váltva vizsgálja az irodalmat: a világháborúk és az 1960-as, 1970-es évek társadalmi mozgalmak. Fanning nagy hatású kutatása több tucatnyi különböző szerző művébe nyújt lenyűgöző betekintést, miközben az olvasók számára a kánon 250 évnyi fejlődésének széles perspektíváját is kínálja. Figyelemreméltó munka. Fanning bibliográfiájában közel 300 elsődleges forrás található.⁹

Fanning kutatása során az 1900 és 1930 közötti időszakban aktív írók nyilvánvaló hiányára figyelt föl, és ennek az irodalmi „recessziónak” a magyarázataként felveti, hogy az 1900 és 1930 közötti időszak – ahogy ő nevezi, egy „nagy méretű kulturális amnézia” korszaka¹⁰ – egy „elveszett generáció” jele. Fanning azt feltételezi, hogy különböző társadalmi erők térítették el az ír-amerikaiakat az ír tapasztalat megírásától, és megjegyzi:

az I. világháború közeledtével az ír-amerikai etnikai öntudatosság kifejezetten visszatetszővé vált számos nem ír-amerikai szemében. Amikor a háború 1914 augusztusában kitört, a britellenes érzelmek erőteljesen az ír-amerikai nemzeti körök ellen fordultak... Az Anglia oldalán tett háborús erőfeszítések és az ír nacionalizmus negatív percepciója a húsvéti felkelést követően együtt vezettek ezekben az években az ír-amerikai öntudat lángjának kihunyásához.¹¹

A számok azonban egy másik történetet mondanak el. Az *1-es ábra* az 1800 és 2005 között megjelent 758 ír-amerikai fikciós mű időbeli eloszlását mutatja.¹²

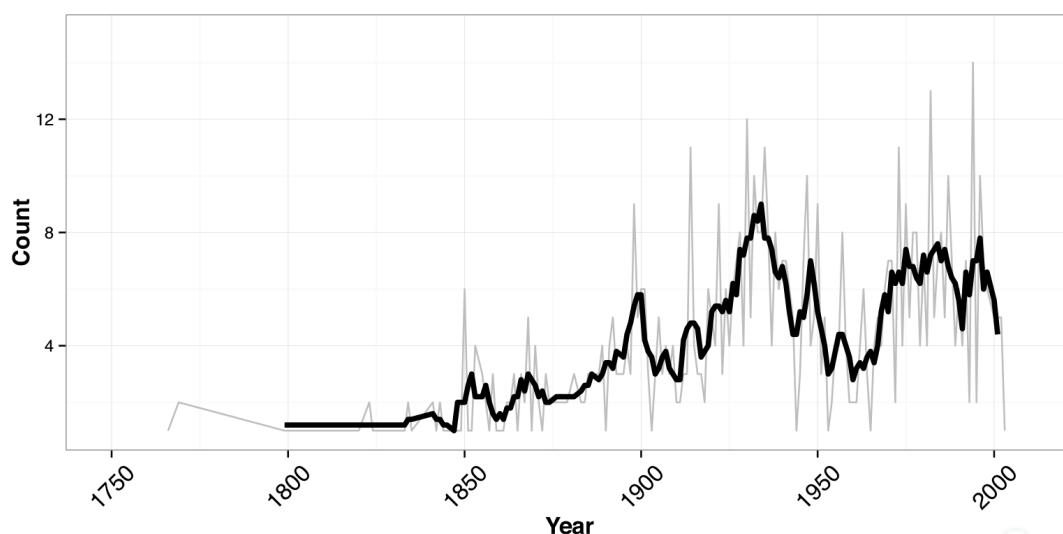
⁸ Vannak más munkák is, amelyek az ír-amerikai irodalomtörténet nagyobb szeletét kísérik meg feldolgozni. Az én munkáim a San Francisco-i öböl területén, Montanában és a Középnnyugaton élő ír-amerikai szerzőkről az újabbak közé tartoznak. M. Jockers, „A Window Facing West: Charles Driscoll’s Kansas Irish,” *New Hibernia Review* 8, 3. sz. (2004): 100–113. M. Jockers, „A Literature of Good Fortune,” in *The Irish in the San Francisco Bay Area: Essays on Good Fortune*, eds. D. Jordan and T. O’Keefe (San Francisco: Irish Literary and Historical Society, 2005), 8–27. M. Jockers, „West of Éire: Butte’s Irish Ethos,” in *All Our Stories Are Here: Critical Perspectives on Montana Literature*, ed. B. Harrison (Lincoln: University of Nebraska Press, 2009), 37–51. Egy másik mű, Ron Ebest *Private Histories* című munkája, amely számos szerzőt tárgyal, akik 1900 és 1930 között voltak aktívak. R. Ebest, *Private Histories: the Writing of Irish-Americans, 1900–1935* (Notre Dame, Ind.: University of Notre Dame Press, 2005).

⁹ Mint olvasó és mint Charles Fanning korábbi tanítványa lekötelezettje vagyok munkájának, nagy hatást gyakorolt rám. Kutatásomat a modern technológia segítette és az a nagyszámú, sokféle mű, amelyeket Fanning kutatott föl, mindez e technikai segítség nélkül a tudományos kitartás mintájaként szolgált.

¹⁰ Fanning, *The Irish Voice in America*, 3.

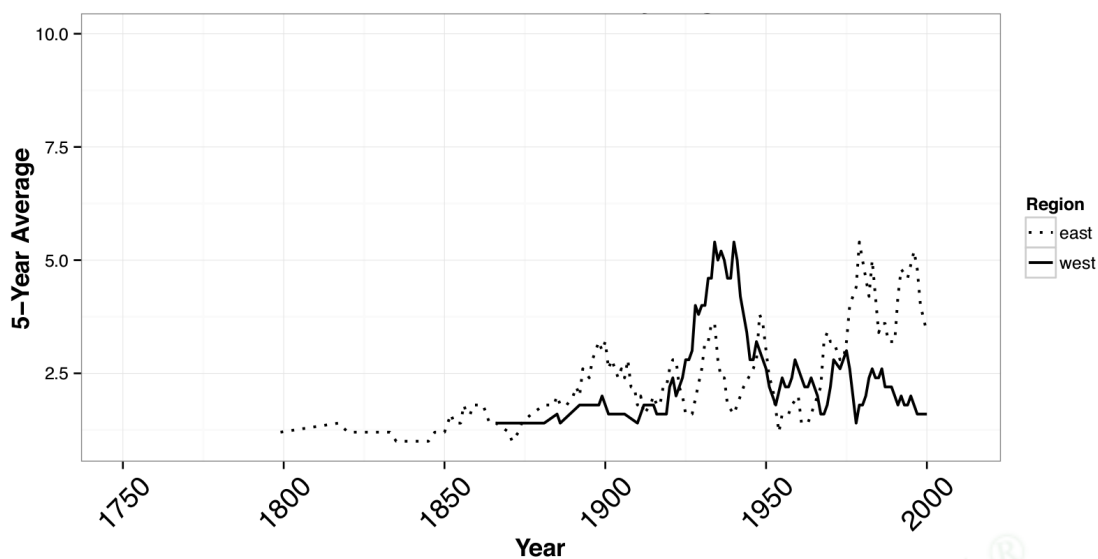
¹¹ Fanning, *The Irish Voice in America*, 238.

¹² Az adatbázis számára 1993-ban kezdtem el gyűjteni a műveket „Az ír-amerikai irodalom a Nyugaton” című tanulmányom részeként. Az adatbázis az évek folyamán nagymértékben bővült, és a Stanfordon tartott ír-amerikai kurzusaimon a hallgatók által végzett kutatás segített kiegészíteni a hiányzó adatokat. A 2000 és 2010 közötti adatok minden táblázatban hiányosak.



1. ábra. Az ír-amerikai fikciós művek időbeli eloszlása

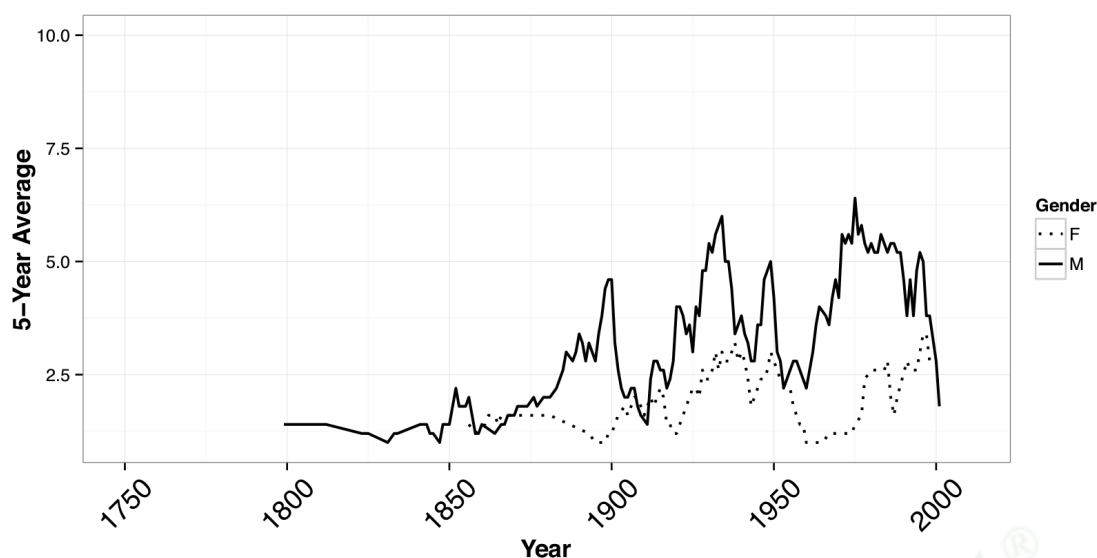
A vékony fekete vonal az adatok ötéves átlagának változását jelzi; a feltűnőbb, pontokból álló az évenkénti aktuális publikációkat. Ez a grafikon tehát pusztán a kiadástörténetet ábrázolja, semmi bonyolultat. Mindazonáltal megjegyzésre méltó, hogy a Fanning feltételezte irodalmi hanyatlás vagy elveszett generáció sokkal rövidebb életű, mint ő elképzelte. Az ír-amerikai publikációk első csúcspontja rögtön a 20. század fordulóján van, amit egy 1910-ig tartó rövid hanyatlás követ. Azután azonban a trend felfelé indul és a publikációk száma pontosan annak a korszaknak a második felében növekszik, amelyet Fanning úgy azonosít, mint amelyben az ír-amerikaiaknak kulturális és társadalmi erők hatására látszólag el kellene némulniuk. Hol van a Fanning által feltételezett elveszett generáció?



2. ábra. Az ír-amerikai fikciós művek időbeli eloszlása régiók szerint

Ha a 2-es ábrán látható módon újrarajzoljuk az eredményeket, azaz a Mississipitől keletre és nyugatra élő szerzők között különbséget teszünk, kezdjük megérteni, mi vezethette Fanningot következtetése levonására: a keleti ír-amerikai szerzők munkáinak kiadása 1900-tól kezdve csakugyan csökken, 1920-ban éri el a mélypontot. A keleti írók az 1930-as évekig nem is kezdik meg a kilábalást ebből a „recesszióból,” és a dolog nem is áll helyre az 1960-as évek végéig, az 1970-es évek elejéig.

Ha csak a pontozott vonalra tekintünk, a keletiekre, a 2-es ábra megerősíti Fanning további észrevételét, mely szerint az ír-amerikai fikciós irodalom a századfordulón virágzott, majd az 1960-as és 1970-es években, amikor etnikai szempontok szerint írni népszerűbb és csábítóbb volt. A nyugatiak vonala azonban egy másik történetet mond el. A nyugati írók némiképp váratlanul tűnnek fel 1900-ban, amivel egy negyvenéves növekedési időszak kezdődik, amely 1941-ben éri el csúcspontját. A nyugati írók egyértelműen uralták a 20. század korai szakaszát.

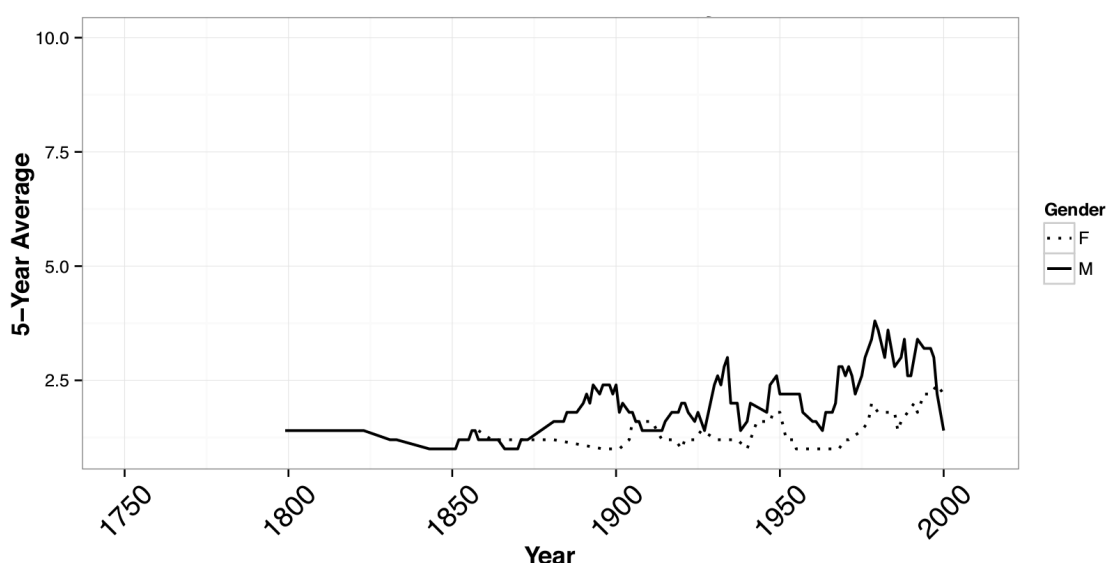


3. ábra. Az ír-amerikai fikciós művek időbeli eloszlása a szerzők neme szerint

A szövegeknek a szerzők neme alapján történő további elkülönítése még többet felfed abból, mi történt ebben az időszakban. Kezdjük a 3-as ábrával, amelyik az összes szöveget a szerzők neme szerint különíti el. Az 1850-es évek kivételével a férfiak irodalmi termelékenysége következetesen nagyobb, mint a nők produktivitása. Az 1850-es évek némiképp különös helyzetétől eltekintve a férfi és a női vonal jellemzően hasonló görbét követ, ami azt valószínűsíti, hogy az ír-amerikai irodalmi alkotások keletkezését a nemi hovatartozáson kívüli erők határozzák meg.¹³

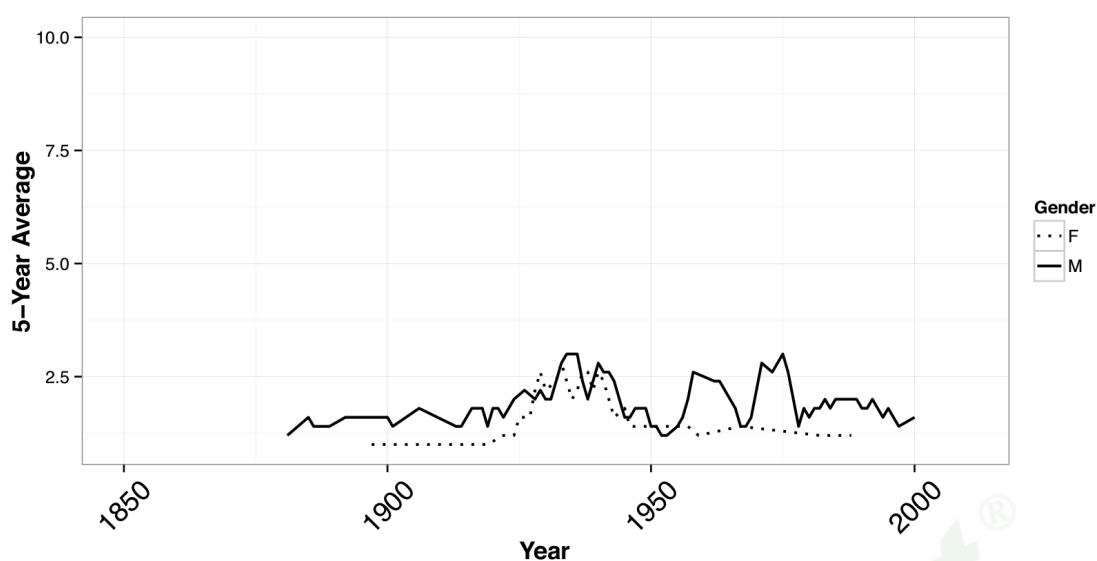
A 4-es ábra ezzel szemben csak a keleti ír-amerikai szövegeket mutatja be a szerzők neme alapján történő bontásban. Úgy látszik, hogy itt, keleten, a férfi szerzők publikációinak száma jóval nagyobb mértékben esett, a századforduló után az évenkénti átlag

¹³ Az adatok korlátozott volta miatt természetesen a vonalak összevetése, még akkor is, ha néhány figyelemre méltó hasonlóságot észlelhetünk. Ha kiszámítjuk a trendeket, azt látjuk, hogy időben előre haladva a termelés lassan emelkedik, és a két trend nagyobb részt párhuzamos, ami az általános növekedés hasonló mértékét jelzi.



4. ábra. Az ír-amerikai fikciós művek időbeli eloszlása a szerzők neme szerint a keleti régióban

2,3-as megjelenés 1916-ra 1,3-as átlagra csökkent. A női publikációk száma 1906 körül lassú, de tartós, tíz éven át tartó emelkedésbe kezdett.



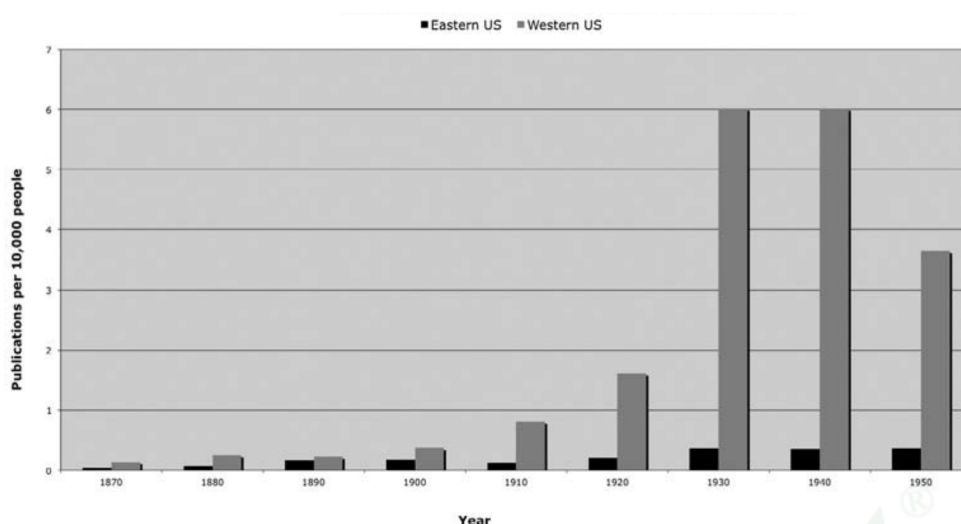
5. ábra. Az ír-amerikai fikciós művek időbeli eloszlása neme szerint a nyugati régióban

Mindazonáltal a keleti publikációk grafikonjánál meglepőbb az 5-ös ábra, amely a nyugati kiadványok adatait tartalmazza a szerzők neme alapján történő bontásban. Míg a nyugati férfi szerzők az 1860-as évektől kezdve az 1930-as évek csúcspontjáig növelik termelékenységüket, az ír-amerikai nők Nyugaton [az Egyesült Államok nyugati partvidéke, a szerk.] csak az 1890-es évek végén tűnnek fel először, és onnantól kezdve számuk gyorsan növekszik, egészen az 1920-as évekig, amikor évenként ugyanannyi könyvet adnak ki, mint nyugati férfi szerzőtársaik. Tekintve, hogy a nők az általános

tendencia szerint kevesebbet publikálnak, ez különösen figyelemre méltó. Ez vagy azt valószínűsíti, hogy a Nyugat valami különlegeset kínál az ír-amerikai nőknek, vagy azt, hogy volt valami rendkívüli azokban az ír nőkben, akik nyugatra mentek, vagy ami még valószínűbb, hogy e kettő kombinációjáról van szó.¹⁴

Egyértelműen úgy tűnik, hogy a nyugati szerzők, férfiak és nők egyaránt, ellenállnak a kelet bármiféle irodalmi recessziójának. Hogy ez annak ellenére (vagy talán éppen amiatt) sikerült-e nekik, hogy a nyugati ír etnikumú populáció jóval kisebb, rendkívül érdekes kérdés. A *6-os ábra* a népszámlálási adatokat is tartalmazza, hogy az ír-amerikai irodalmi termést a keleti és nyugati demográfia kontextusában vizsgálhassuk. A táblázat a térség tízezer ír születésű bevándorlójára eső ír-amerikai könyvek számát mutatja. Természetes feltételezés lenne, hogy pozitív összefüggésnek kell lennie a populáció mérete és a populációban előforduló lehetséges írók számát illetően. Az adatok azonban éppen az ellenkezőjét tárják fel: a ritkábban lakott Nyugat egy főre eső könyvtermése nagyobb.

Az irodalmi termelékenység ellenkező trendjét feltárva ezek az adatok még inkább problémássá teszik Fanning állítását, mely szerint az ír-amerikai szerzők a 20. század elején elhallgattak. A keleti kiadói központoktól földrajzilag messzebbre, az ír-amerikaiak a Nyugaton eltávolodtak az ír kultúra elsődleges elosztóhelyeiként szolgáló városoktól is, Bostontól, New Yorktól és Chicagótól. Mivel nem éri őket a hazai kultúra folyamatos hatásához, azt váránk, hogy a Nyugat ír bevándorlói gyorsabban alkalmazkodnak a befogadó kultúra gyakorlataihoz és tendenciáihoz. Még azt is elképzelhetnénk, hogy az ír-amerikai kultúra támogatásának hiányában az ír származású amerikai írók eltávolodnak az etnikai identifikációktól, és kevesebb etnikailag meghatározott fikciós szöveget írnak. Ám épp ellenkezőleg, a nyugati ír-amerikaiak egy főre számítva exponenciálisan nagyobb arányban írnak arról, hogy milyen írnek lenni Amerikában, mint keleti társaik.



6. ábra. Az ír-amerikai fikciós művek bevándorlók száma szerinti eloszlása

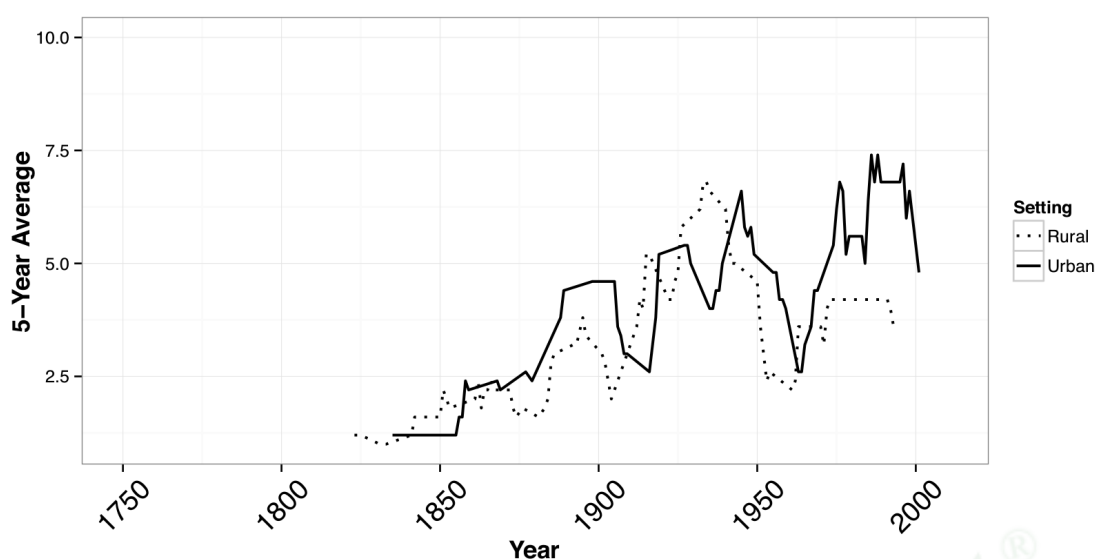
¹⁴ Másutt amellet érveltem, hogy a két faktor kombinációja a legvalószínűbb. Jockers, „A Literature of Good Fortune,” 8–27.

A 6-os ábra oszlopai a 10000 ír bevándorlóra jutó könyvek számát mutatja. A szürke oszlop a nyugati szerzők termelékenységét mutatja, a fekete a keletiekét. Az 1930-ban például 10000 ír bevándorlóra 6 nyugati író által publikált mű esik, ugyanebben az évben a Mississippitől keletre az ír-amerikai szerzőktől tízezer főre 0,3 könyv jut, azaz 30000 ír bevándorlóra jut egy könyv.¹⁵ Az 1900–1940 közötti időszakon kívüli évekre is igaz, hogy a nyugati ír-amerikai szerzők keleti társaiknál többet publikálnak az egy ír bevándorlóra jutó könyvek számát tekintve. Jóllehet igaz, hogy 758 szöveg 250 évre elosztva nem egy jelentős adatmennyiség, az adatok évtizedenkénti csoportosítása és az adatok tízéves átlagának vizualizációja hasznos annak eldöntésében, lényegében melyek a látens makrotrendek. Az adatok gyér voltára tekintettel nem lehet erőteljes következtetést levonni, de az elemzés mégis tanulságos.¹⁶ Azt valószínűsíti, hogy a földrajz vagy a kultúra valamely, a földrajzzal kapcsolatos jelensége hatással van az ír-amerikai irodalom termelékenységére.

A fikciós helyszínek alapján csoportosított szövegeket vizsgálva hasonló utat járhatunk be. A szerzőhöz tartozó metaadatok mellett – nem, lakhely, földrajzi régió – az adatbázisban található munkák fikciós helyszínei is kódolva vannak. Minden egyes tétel mellett ott áll, hogy „vidéki”, „városi” vagy „vegyes”. Csak azokat a munkákat ábrázolva, amelyek egyértelműen vagy a vidéki, vagy a városi kategóriába esnek, további információkat nyerünk az ír-amerikai irodalom eloszlásának és szerkezetének megértéséhez. A 7-es ábra azon 1760 és 2010 között megjelent fikciós munkákat tartalmazza, amelyek vagy vidéki, vagy városi környezetben játszódnak. Kifejezetten az 1900 és 1940 közötti időszakot vizsgálva az 1930-as években a városi munkák dominanciája felé vezető trend rajzolódik ki, amely 1940 körül éri el csúcspontját. Ami nem látszik ebből, az az, hogy valójában a nyugati szerzők írták a városi helyszínen játszódó művek túlnyomó részét. 1900 és 1930 között a nyugati ír-amerikai szerzők helyezték fel a városi Amerikát a térképre, ami arra indít, hogy Fanning megfigyelését az elveszett nemzedékről tovább vizsgáljuk. Úgy tűnik, hogy ez az elveszett nemzedék valójában egy keleti, valószínűleg férfiakból álló ír-amerikai elveszett nemzedék volna, amely városi témákról írta.

¹⁵ A népszámlálási adatok csak az első generációs ír bevándorlókról szolgáltatnak információt, míg a publikációs számokat a generációfaktor nem befolyásolja. Ennélfogva e könyvek szerzői lehetnek akár elsőgenerációs ír bevándorlók, akár ír bevándorlók második vagy harmadik generációs leszármazottai. Az egyik feltételezés, amit az adatok megcáfolnak, hogy az új bevándorlók állandó beáramlása állandósítja az etnikai és kulturális témák iránti érdeklődést; minél több új bevándorló, annál inkább feltételezhető az ír témák irodalmi feldolgozása. Épp ellenkezőleg, a bizonyíték azt valószínűsíti, hogy nincs összefüggés a bevándorló népesség és az etnikai orientációjú fikciós irodalom között. A népszámlálási adatok a történeti Census Browserből származnak, 2007. január 29-én letöltve az University of Virginia Geospatial and Statistical Data Centeréből: <http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html>.

¹⁶ 758 szöveg 250 évre elosztva átlagosan mindössze három szöveget jelent, és a valóságban a korai években sokkal kevesebb volt. Ahhoz, hogy következtetéseket vonjunk le, egy szélesebb és egyenletesebben eloszló szövegtörzshöz volna szükség. Mindazonáltal úgy sejttem, hogy ez a – szövegek „aktuális” sokaságának értelmében – teljes korpusz megközelíti azt a mennyiséget, amely e kategóriában egyáltalán létezhet. Ebben az értelemben az eredmények meggyőzők.



7. ábra. Az ír-amerikai fikciós művek eloszlása az ábrázolás színhelye szerint

Nehéz nem levonni ezt a következtetést, különösen, ha meggondoljuk, hogy az amerikai írek történelme és még inkább az ír-amerikai irodalom története kétségtelenül keleti dominanciájú. Casey és Rhodes 1979-es gyűjteményének, az *Irish-American Fiction: Essays in Criticism* előszavából is látható, mennyire elterjedt a kelet dominanciája. William V. Shannon az előszóban arról ír, hogy az esszégyűjtemény „az amerikai-ír irodalom teljes egészét” vizsgálja, valamint „bemutatja az ír közösség Egyesült Államokban szerzett széleskörű, változatos tapasztalatait.”¹⁷ Jóllehet Shannon lelkesedése a gyűjtemény meghatározó jelentőségének köszönhető, állítása teljes mértékben félrevezető. A gyűjteményben közölt esszék vizsgálata feltárja, hogy korántsem teljes körű elemzésről van szó, hanem sokkal inkább szinte kizárólag keleti, városi környezetben élő ír-amerikai szerzők elemzéséről. A Nagy-síkság és az attól még nyugatabbra lévő területek írói teljesen hiányoznak, ennél fogva a gyűjtemény 1806692 négyzetmérföld ír-amerikai irodalmi tapasztalatának „teljes egészét” kihagyja.¹⁸

Ez a gondatlanság, tekintve az amerikai és ír történészeknek a Nyugaton élő írek felé forduló kritikai figyelmét, meglepő. Jóllehet igaz az, hogy a történész tanulmányok többsége a keleti, városi környezetben élő ír közösségekkel foglalkozik, mint Boston, New York, de legalábbis az 1970-es évek óta jelentős erőfeszítések történtek a Mississippin túl élő írek tapasztalatainak leírására is. 1977-ben Patrick Blessing arra mutatott rá, hogy az írek Amerikájáról a történészek által alkotott aktuális kép „túlságosan homályos”, mert eltúlzottan olyan városokra koncentrál, mint Boston.¹⁹ Akkortájt, jegyzi meg, „az Egyesült Államokban élő írekkel foglalkozó tanulmányok a keleti part városi, ipari központjaira korlátozódnak.”²⁰ Az amerikai írekről alkotott

¹⁷ William V. Shannon, „Preface,” in *Irish-American Fiction: Essays in Criticism*, eds. D. J. Casey and R. E. Rhodes (New York: AMS, 1979), ix.

¹⁸ A Mississippitől nyugatra lévő területek becsült mérete.

¹⁹ P. Blessing, „West among Strangers: Irish Migration to California, 1850 to 1880” (Ph.D. diss., University of California, Los Angeles, 1977), 2.

²⁰ Blessing, „West among Strangers,” 2.

tudományos vélemények tehát szintén túlságosan keletközpontúak voltak, és nagyrészt az éhségkorszak bevándorlóinak küzdelmeire összpontosultak. Mindazonáltal Blessing tanulmánya a Kaliforniába történő ír bevándorlásról a fennálló elképzelések újragondolására készítetett. A Blessing által idézett 19. század végi népszámlálási adatok rámutatnak, milyen kevesen ismerték fel: nemcsak nagy számban éltek írek Nyugaton, hanem néhány esetben ők alkották a domináns etnikai csoportot. 1850 és 1880 között például az írek alkották a legnagyobb bevándorló közösséget Minnesotában, Coloradóban, Új-Mexikóban, Wyomingban, Montanában, Arizonában, Nevadában és Washingtonban, és a második legnagyobb közösséget alkották Utahban, Idahóban, Oregonban és a Dakota-vidékeken.²¹ Ezeket az adatokat és más kvalitatív bizonyítékokat is felhasználva Blessing amellett érvelt, hogy az a régóta hangoztatott vélemény, mely szerint „az ír bevándorlók elutasították, hogy a városi észak-keletről elköltözzenek”, komoly újragondolást igényel.²² Még fontosabb, hogy megcáfolta azt az elképzelést, mely szerint az íreknek „averzióik voltak a Nyugatra költözéssel szemben,” rámutatva, hogy 1850-ben „200000 ír élt a keleti államokon kívül.” 1880-ra a szám „majdnem 700000-re nőtt, az Egyesült Államokban élő teljes csoport 36%-ára.”²³ Erről a kelet-központúságról írja Blessing:

Az, hogy a történészeknek foglalkozniuk kell a keleti városokban csoportosuló írekkel, érthető, többségük évtizedekkel az Egyesült Államokba való érkezés után is a belépési pont közelében található... Azt hangsúlyozni azonban, hogy az írek elutasították, hogy az urbánus észak-keletről eslköltözzenek, figyelmen kívül hagyja az Emerald szigetről érkező nagyarányú beköltözőket, akik elhagyták a régiót.²⁴

Ha a Nyugat elhanyagolása probléma, még inkább az figyelmen kívül hagyni az ír-amerikai nőket. Az ő figyelmen kívül hagyásukról írva jegyzi meg Caledonia Kearns, hogy „az ír-amerikai nők történetét ebben az országban Hasia Diner alapvető könyve, az *Erin's Daughters in America* megjelenéséig alig észlelték.”²⁵ Ha a nők által művelt ír-amerikai irodalomra kerül sor, a helyzet még rosszabb. Kearns azzal folytatja, hogy rámutat, nagyon „kevés tanulmány [született] az ír-amerikai szépirodalomról [...] és ha [az ír-amerikai szépirodalomról] van szó, a férfi írókra esik a figyelem.”²⁶

Feltéve, hogy a tudósok, az olvasóközönség és a könyvkritikusok tényleg bűnösök abban, amit Blessing és Kearns sugall, tanulságos meggondolni, mi vethető ezen előítéletek ellen az ír-amerikai irodalmi kánonon belül. Egy olvasó, egy kritikus vagy egy tudós, aki keleti, városi férfiak iránti elfogultsággal rendelkezik, kevés művet talált az ír-amerikai korpuszban 1900 és 1940 között. E negyvenéves periódus mindössze 5%-a felel meg annak a kritériumnak, hogy keleti férfi írta és városi környezetben játszódik. Ha az érdeklődést teljesen megfordítjuk (nyugati nő és vidéki környezet), majdnem

²¹ Blessing, „West among Strangers,” 169.

²² Blessing, „West among Strangers,” 162.

²³ Blessing, „West among Strangers,” 163.

²⁴ Blessing, „West among Strangers,” 163.

²⁵ C. Kearns, *Cabbage and Bones: An Anthology of Irish American Women's Fiction* (New York: Henry Holt, 1997), xvii.

²⁶ C. Kearns, *Cabbage and Bones*, xviii.

kétszer annyi könyvet találunk. Visszatekintve látszik, hogy Fanning átfogó munkája az ír-amerikai irodalomról alapvetően kanonikus és anekdotikus. Mindazonáltal ezért nem hibáztatható, hiszen sem az adatok, sem a módszer nem állt rendelkezésére, hogy másképp csinálhassa. Olyan sok anyag van, amit a szoros olvasással (*close reading*) és a tudományos szintézis hagyományos módszereivel áttekinthetünk.²⁷

Az itt alkalmazott makroanalitikus perspektíva hasznos, jóllehet hiányos helyesbítést jelent. A megközelítés azonosítja a nagyobb trendeket, valamint lehetővé teszi a korpusz növekedési és hanyatlási periódusainak kritikai megfigyelését. A grafikonok és az ábrák 758 fikciós mű korpusza alapján készültek, és jóllehet nem értelmetlen vagy éppen lehetetlen elvárni, hogy egyetlen tudós elolvasson és kibányásszon ennyi szöveget, ez a feladat meglehetősen sok időbe kerülne, és az emberi szintéziskészítés határait feszegetné. Mindazonáltal még a legfelkészültebb hagyományos tudós számára sem mindennapi feladat, hogy több mint 750 szöveg szoros olvasása nyomán támadt különböző benyomásait koherens és kezelhető irodalomtörténetté sűrítse. Bizonyosan nem lehetetlen, de biztos, hogy nem praktikus. A bizonyítékok közötti szemezgetés – a nagy hipotézis támogatása érdekében – elkerülhetetlennek tűnik a szoros olvasásra épülő hagyományos tudományos feldolgozásban.

Mivel mindent elolvasni, nem kifejezetten praktikus, a hagyományos irodalomtörténészek között az a tendencia, hogy egy irodalmi korszakról korlátozott szövegminta alapján vonják le következtetéseiket. Az egyedi általánosításának gyakorlata akkor különösen veszélyes lehet, amennyiben a vizsgált szöveg nem reprezentatív az egészre nézve. Érdeemes megismételni, ami korábban már idéztünk. A 19. századi brit regényről beszélve Moretti a következőket írja:

Mindannyian az irodalmi mező kicsiny szeletén dolgozunk: például egy kétszáz regényből álló kánon elég nagynak tűnik a 19. századi Nagy-Britanniára nézve [...] de ez még mindig kevesebb, mint a ténylegesen publikált regények egy százaléka [... Egy] ilyen méretű mezőt nem érthetünk meg, ha egyedi esetekről szerzett ismereteinket egybefűzzük.²⁸

Az ehhez hasonló kommentárok miatt néhányan azzal váddal lépnek föl, hogy Moretti vállalkozása szembemegy az irodalom valódi tanulmányozásával, és talán még veszélyezteti is azt. William Deresiewicz a Moretti által szerkesztett *The Novel* című tanulmánygyűjteményről írott kritikájában Morettit egyféle irodalmi „hódítóként”

²⁷ Fontos, hogy rámutassunk, Fanning munkájának elsődleges célja az *Irish Voice in Americában* az, hogy felhívja a figyelmet azokra a munkákra, amelyek leginkább reprezentálják az egyesült államokbeli ír tapasztalatot. Fanning nem egyszerűen azokat a munkákat akarta kiemelni, amelyek a legpontosabban festik le az írnek lenni Amerikában kérdését, hanem azokat a munkákat – és ez még fontosabb –, amelyek a legnagyobb erővel, irodalmi stílussal festik le az ír tapasztalatot. Fanning sikere döntő abban, hogy a hagyomány legjobb íróira irányította figyelmünket, az a kérdés viszont nyitva marad, hogy ezek az új nyugati írók esztétikailag összemérhetők-e keleti társaikkal – ezt egy másik projektben fogom feltenni. A Fanning által említett elveszett generáció valójában a jó írók elveszett generációja. Itt nem célom, hogy elkülönítsem a könyveket az általam érzékelt esztétikai érdemeik alapján.

²⁸ F. Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History* (London and New York: Verso, 2005), 4.

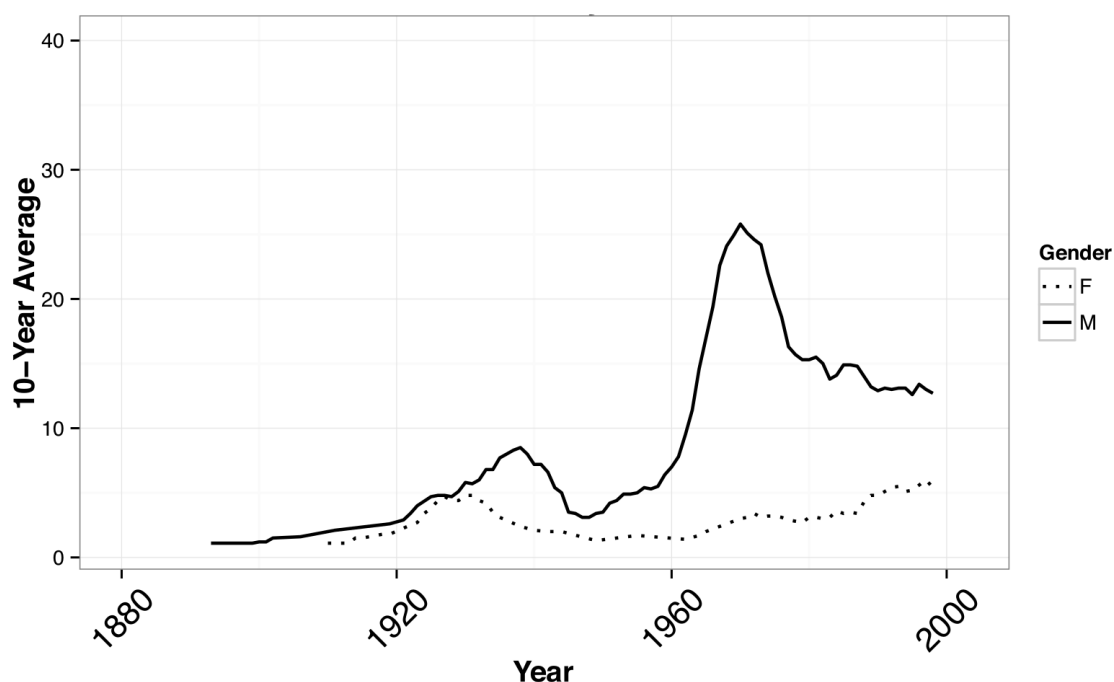
mutatja be.²⁹ Úgy ábrázolja Morettit, mint egy irodalmi-kritikai „harc” katonáját. Deresiewicz egyértelművé tette, hogy ellenáll, és úgy látja, hogy az irodalom tanulmányozása során nincs hely mindkét – a makro (távoli) és a mikro (szoros) – megközelítés számára. Az ő tudományos harcterén a tudósnak vagy az egyik, vagy a másik módszert kell választania. Deresiewicz azonban téved. Moretti intenciója nem az, hogy teljesen megsemmisítse a hagyományos irodalomtudományt a távoli olvasás sorozatvetőjének alkalmazásával, a makroanalízis nem a szoros olvasás ellen bevetett versenytárs. Mind az elmélet, mind a metodológia célja bizonyítékok felfedezése, bizonyítékok szolgáltatása. Ez a bizonyíték különbözik attól, amit szoros olvasás révén nyerünk, de ez is bizonyíték, fontos bizonyíték. Időnként az új bizonyíték megerősíti mindazt, amit anekdotikus tanulmányokban már összegyűjtöttünk (mint Fanning megfigyelését az ír-amerikai irodalom 1960-as, 1970-es évekbeli virágzásáról). Más alkalommal a bizonyíték megváltoztatja elképzelésünket arról, amit tudni véltünk. Az eredmény mindkét esetben ugyanaz: pontosabb kép tárgyunkról. Ez nem a tudományosság hagyományos módszereinek „meghódítására” és teljes eltörlésére irányuló radikális harc vagy egyéni erőfeszítés.

* * *

A szövegeknek a metaadatok alapján történő egyszerű megszámlálása, elrendezése az ír-amerikai irodalom fenti elemzéséhez adatokat biztosít, de nem ad elegendő anyagot egy átfogó tanulmányhoz. Bármely mérőszám alkalmas arra, megkíséreljünk általa jobban megérteni egy irodalmi korszakot vagy kánont. Érdeklődhetünk pusztán az irodalmi termés iránt, vagy az iránt, a termés hogyan viszonyul a népességhez. Az „egy főre eső könyvek száma” típusú elemzés, mint amilyen a 6-os ábrán látható, hasznosnak bizonyulhat az olyan állítások érvényességnek megbecsülésére, amelyek gyakran elhangzanak az irodalmi „reneszánszok” kapcsán. Az 1930-as és 1940-es évek afro-amerikai drámájának (*black drama*) virágzása néhány kulcsfontosságú író kreatív géniuszának tulajdonítható-e, vagy az oly nagy mennyiségű mű egyszerűen a potenciális írók számából következő valószínűség?

Annak érdekében, hogy a kérdést megbízható módon megbecsülhessük, az a helyes, ha először az egy évre jutó drámák puszta számát tárjuk fel. Ezután kiszámolhatjuk a drámák évtizedenkénti vagy valamely adott időintervallumra eső átlagát. Továbbá helyes volna, ha összevetnénk az afro-amerikai drámák termését általában az amerikai dráma szélesebb kánonával. Szignifikánsan nagyobb volt az afro-amerikai drámaírók termése a fehér szerzőkénél? Ugyanebben az időszakban az afro-amerikai közösségen kívül is megfigyelhető a dráma virágzása? Nyilvánvaló, hogy az e kérdésekre adott válaszok jelentősen befolyásolják, hogyan értelmezzük az amerikai drámát általában, illetve magát az afro-amerikai irodalmat. A kérdésekre adott válaszok nemcsak a darabok hozzáférhetőségétől függenek, hanem a hozzáadott információktól is, amelyeket a darabokról szóló metaadatok nyújtanak. Azok az olvasók, akiket érdekel e kérdések vizsgálata, ilyen metaadatokat (és még többfélét is) találnak az Alexander Street Press *Black Drama* című, gyönyörűen összeállított gyűjteményében. A jelenlegi

²⁹ W. Deresiewicz, „Representative Fictions,” *Nation*, 2006. dec. 4.



8. ábra. Az afro-amerikai drámák időbeli eloszlása nemek szerint

(második) kiadás 1450, a 18. század közepétől napjainkig írott drámát tartalmaz Észak-Amerikából, az angol nyelvű Afrikából, a Karib-szigetektől és más afrikai diaszpórával rendelkező országokból, több mint 250 drámaírótól.³⁰ A 8-as ábra a *Black Drama* darabjainak nemek szerinti kronológiai eloszlását mutatva, valószínűleg többet mond az Alexander Street Press összeállításának gyakorlatáról, mint az afro-amerikai dráma történetéről.

Bárhogyan legyen is, a táblázatból ugyanaz derül ki. Azt sugallhatja például, hogy a fekete férfiak nagyobb valószínűséggel drámaírók, vagy talán volt valami különleges az 1970-es évek körül, ami a női szerzőket nem, viszont a férfi szerzőket arra ösztönözte, hogy drámát írjanak. Feltárhatja, hogy az 1920-as, 1930-as évek ún. Harlem Renaissance-a olyan időszak volt, amelyben a női drámaírók vagy aktívabbak voltak, vagy az utókor jobban értékelte őket. Fölösleges is mondani, hogy ezek olyan kérdések, amelyeket azoknak kell feltenniük, akik jártasak az afro-amerikai irodalom területén. Az olyan gyűjtemények, mint ez az Alexander Street Presstől, termékeny alapul szolgálhatnak a mélyebb makroanalitikus kutatásoknak, mivel a gyűjtemény gazdag metaadatban és teljes szöveget is közöl.

Láttuk, hogy a földrajz és az ír-amerikai termelékenység között szoros kapcsolatot tételezhetünk fel. A pusztán termelékenység megértése mellett az ír-amerikai adatbázisban tárolt metaadatok korlátozott módon betekintést engednek a regények tartalmába és stílusába is. Az Egyesült Államokban megszokott eljárás az irodalom földrajzi szempontú feltárása, műfajok földrajzi különbségei alapján történő elkülönítése, megkülönböztetése, osztályozása: a New England-i költők, a déli regionalista

³⁰ Lásd a <http://bld2.alexanderstreet.com/> oldalt (előfizetést igényel)!

írók, a nyugat-amerikai irodalom stb. Az ír-amerikai adatbázis korlátozott, végső soron nem más, mint egy kibővített bibliográfia. Ugyanakkor egy hasonlóan gazdag metaadatokkal rendelkező teljes szövegű adatbázison olyan lekérdezéseket is végre tudunk hajtani, amelyek lehetővé teszik annak feltárását, hogy az irodalmi stílus és a nyelvi kifejezések közötti szerzői választások miként változnak a különböző régiókban. Ahogyan a nyelvészek a dialektusokat vagy regisztereket földrajzi kontextusban tárgyalják, a földrajzi kontextussal összekapcsolt „irodalmi” stílus tanulmányozása azt a lehetőséget nyújtja, hogy feltárjuk, a nyugati írók bizonyos stílusjellemzőkkel bírtak, amelyekhez hasonló kifejezést nem találunk a keletieknél. Amíg a nyelvészeket a nyelv általános fejlődése érdekli, az irodalmi stílus kutatóit az érdekli, hogyan válik az író technikája vagy az előadásmód egyedivé. A déli regionalistáknak van elkülönült írói stílusa vagy dialektusa, amely megkülönbözteti őket északi társaiktól? A déli írók prózájának „ritmusa” lassabb? A déli írók hajlamosabbak a metaforikus kifejezésekre? Az északiak nagyobb szótárral dolgoznak? Természetesen találnánk lexikai különbségeket a déli és az északi írók között, a regionális írók elkerülhetetlenül a helyi térre, szokásokra utalnak. Ám érdekes kérdés, vajon létezik-e észlelhető stilisztikai különbség ezeken a toponímiákon, regionális különbségeken túl. Még nehezebb kérdés annak a vizsgálata, hogy a regény műfaj milyen mértékben határozza meg a regénystílust, vagy hogy az irodalom milyen mértékben működik megjósolható stilisztikai kifejezésekből alkotott rendszerként. E nagyobb kérdések a könyv későbbi fejezeteire maradnak. A makroanalitikus módszertanba történő további bevezetésre egy kevésbé ambiciózus kérdés bizonyul hasznosnak: az ír-amerikai irodalomban milyen mértékben gyakorol hatást a földrajz a tartalomra? Teljes szövegű korpusz nélkül valamivel jobban kell fókuszálni a kérdést: a nyugati ír-amerikai írók több vagy kevesebb valószínűséggel határozzák meg etnikai hátterüket, mint a keleti írók? A maga gyakorlatiasságában nemcsak jól kezelhető, de fontos kérdés is ez, mivel az ír-amerikai írók jó része, részben néhány nagyon prominens író, köztük Fitzgerald és O'Hara szándékosan kerülte az etnikai markerek használatát, amelyek műveiket „írként” bélyegeznék meg. Az ír-amerikaiak egészen John F. Kennedy megválasztásáig nem hagyták teljesen maguk mögött (akár ténylegesen, akár csak szimbolikusan) a katolikus- és írelles előítéleteket, amelyek az 1800-as évek végét jellemezték, amikor a szegény ír katolikus bevándorlók elárasztották az amerikai partokat. Kétségtelen, néhány ugyanilyen stigma, sztereotípiá máig fennmaradt, de egészében véve ír-amerikainak nevezni valakit többé már nem becsmérő. Mindazonáltal azon ír-amerikaiaknak, akik írtak, és íróként sikert akartak elérni a nagy éhínséget követően, egy ír vezetéknev nem volt szükségszerűen értékesíthető előny. Amennyiben egy író nem akarta kihasználni az ún. írséggel kapcsolatos szokásos sztereotípiákat, legjobban tette, ha teljesen elkerülte az ír témákat és karaktereket. Ha elfogadjuk ezt a premisszát, akkor azok, akik etnikai elemekkel kötelezik el magukat szövegeikben, érdemesek arra, hogy az amerikai szerzők külön alcsoportjaként vizsgáljuk őket.

Claude Duchet azt írta, hogy a regények címei „kódolt üzenetek – egy piaci szituációban.”³¹ Valóban, a regények címei – még ha nem is bírnak megfelelő elmélettel – hihetetlenül hasznos és fontos aspektusai az irodalmi adatnak. Amint azt Moretti

³¹ F. Moretti, „Style, Inc.: Reflections on Seven Thousand Titles (British Novels, 1740–1850),” 134.

igazolta, teljesszöveges archívum hiányában, még ha tökéletlen módon is, a művek címei hasznos képviselői maguknak a regényeknek. Az ír-amerikai irodalom esetében ez kiemelten fontos lehet, ti. az ír témákkal szembeni történelmi vagy a kulturális háttérből fakadó intolerancia miatt a művek címei különleges súllyal bírnak, mivel a cím az, ahol egy szerző először találkozik olvasójával. Milyen mértékben azonosítják tehát magukat az ír-amerikai szerzők írként könyveik címében? Bizonyos szótípusok – jelen esetben etnikai és vallási jelölők – időbeli, régiónkénti gyakoriságának és eloszlásának a vizsgálata a kiadástörténet korábban bemutatott elemzését ismételtén demisztifikálja.³² Az ír-amerikai irodalmi alkotások címében az etnikai és vallási jelölők használatának első hulláma az 1850 és 1870 közé eső évtizedekre esik: az etnikai vagy a katolikus jelölők (vagy mindkettő) a címek 30%-ában tűnnek fel az 1850-es években, 33%-ban az 1860-asokban és 43%-ban az 1870-esekben. Ezek az írószázi nagy éhínséget követő közvetlen évek. Az e kifejezéseket alkalmazó szerzők közül sokan bevándorlók: első generációs ír-amerikaiak arról írnak, amit a legjobban ismernek, szülőföldjükről, népükről. E szerzők nagy része tanult ember, akik osztályuknak, vagyonuknak köszönhetően Amerikába érkezve elkerülhetik a közvetlen diszkriminációt. Az 1880-as években azonban pontosan nullaszer fordulnak elő az etnikai jelölők az ír-amerikai szövegek címében, és 1890-től 2000-ig a jelölőszavak csak az 1890-es években érik el a 11%-ot, de aztán soha nem haladják meg a 6%-ot. Ez természetesen felvet egy kérdést: mi történt 1880-ban, 1880 körül, ami a meredek csökkenést okozta? Most az újonnan megtalált tényektől az értelmezés felé kell fordulnunk.

A legnyilvánvalóbb válasz a bevándorlást követő diszkrimináció. 1880-ra az éhínség elől menekülő bevándorlás legnagyobb része lezajlott, és az Egyesült Államokban valószínűleg volt elegendő ír- és katolikusellenes érzület ahhoz, hogy az ír identitást vállalni kényelmetlen és nem túl jövedelmező elképzelésnek tűnt.³³ Az adatok részletesebb vizsgálata azonban feltárja, hogy 1880 után ezek a jelölők egyedül a Mississippi-től keletre élő írók által, az ott élő írekről írott könyvek címében csökkentek. A Mississippi-től nyugatra élő ír szerzők továbbra is alkalmazzák irodalmi alkotásaik címében ezeket a jelölőket, különösen az 1890 és 1900 közötti évtizedben, amikor minden egyes szöveg használja az etnikai markereket, jelölőket. Ebben az évtizedben az összes kiadott ír-amerikai szöveg 11%-a származik a Mississippi-től nyugatra élő szerzőktől.

Az etnikai jelölőket alkalmazó írók túlnyomórészt vidéki helyszínekről írnak. 1900 után csak néhány jelölő tűnik föl a vidéki környezetben játszódó művek címében, mint például Charles Driscoll könyvében, a *Kansas Irishben*.³⁴ Amint az 1-es táblázatban láthatjuk, 1900 és 1909 között a vidéki környezetben játszódó könyvek címének 7%-ában van egyértelmű etnikai azonosító, 1910 és 1919 között 22%-ában. Az ugyanazokban az

³² Az „etnikai jelölő” azokat a szavakat tartalmazza, amelyeket egy tipikus olvasó írként azonosít, vagy az ír etnikumra asszociál róluk. E jelölők közé tartozik néhány nyilvánvaló szó, mint az Irish, Ireland, Erin, Dublin, Shamrock és a Donegal, valamint néhány kevésbé nyilvánvaló, de ugyanennyire telített szó és név, mint többek között a priest, Patrick, parish, diocese, Catholic, Lonigan, Murphy, O’Phelan, O’Regan, O’Neill, O’Mahony, O’Donnells, O’Flarrity, O’Halloran, O’Shaughnessy.

³³ Ez Fanningnak a korai 1900-as évek kontextusában megalkotott magyarázatához hasonló értelmezés.

³⁴ Driscoll regényéhez írt kritikai bevezetőm a *Kansas Irish* legújabb kiadásában olvasható, amely a wichitai Rowfant Pressnél Kansasban 2011-ben jelent meg.

évtizedekben városi környezetben játszódó, etnikai azonosítót alkalmazó művek csak 1940 és 1949 között tűnnek fel 7%-ban, 1970 és 1979 között csak 4%-ban, 1980 és 1989 között mindössze 2%-ban. Az itt feltűnő minta meglehetősen egyértelmű: a nyugati írók és a vidéki környezetben játszódó művek írói könyveik címében nagyobb valószínűséggel jelölik az írséget, nagyobb valószínűséggel jelzik előre leendő olvasóknak ezt a fajta érdeklődést.³⁵ A nyugati ír-amerikai írók nyilvánvalóan nagyobb magabiztossággal vagy legalábbis nagyobb érdeklődéssel nyilatkoznak tárgyuk írségéről.

Nyilvánvaló, hogy a cím a teljes könyv reklámja, vagyis a marketing része, de gyakran azt a célt is szolgálja, hogy sűrített módon kifejezze a szerzőnek saját munkájáról alkotott benyomását. Gondoljunk például a két következő címre: *The Irish Emigrant: An Historical Tale Founded on Fact* és *The Aleins*. Mindkettő az Egyesült Államokba történő ír bevándorlással foglalkozik. Az első egy 1817-ben kiadott könyv, ez egyértelműen azonosítja az ír bevándorlás témáját. A másik könyv 1886-ból ugyanezzel a témával foglalkozik, de nem teszi lehetővé, hogy etnikai kifejezésekkel határozzák meg. Lehetséges, hogy a korábbi szerzője (az éhínség előtt, az amerikai társadalmat átható, éhínség utáni ír ellenes előítéletek előtt) magabiztos volt tárgya ír voltának azonosításakor? Pusztá sejtés, természetesen, de egy efféle hipotézisnek van intuitív értelme.

Decade	Rural texts (%)	Urban texts (%)
1900–1909	7	0
1910–19	22	0
1920–29	0	0
1930–39	4	0
1940–49	0	7
1950–59	0	0
1960–69	0	0
1970–79	18	4
1980–89	29	2
1990–99	0	0
2000–2005	9	0

1. táblázat. Vidéki és városi környezetben játszódó, etnikai jelölőket használó művek százalékaránya

Az ír-amerikai irodalommal és kultúrával foglalkozó tudósok számára az itt feltárt földrajzi és kronológiai elemzés megerősíti azt a fontos gyanút, hogy a kultúra és a földrajz nemcsak az írói hajlamot határozza meg, hanem a szöveg tartalmát és

³⁵ Felmerült bennem annak a lehetősége is, hogy ezt a címadási mintázatot a kiadók befolyásolják és nem a szerzők. A kiadót és a kiadási helyet a metaadatok szintén rögzítik, és nem találtam semmit ezekben az adatokban, ami azt sugallta volna, hogy a szerzői döntésen kívül bármi más befolyásolta volna a trendeket.

érzelmi viszonyulását is. A történészek, köztük James Walsh, David Emmons, Patrick Blessing és Patrick Dowling már dokumentálták,³⁶ hogy az amerikai nyugatra tartó ír bevándorlóknak eltérő kihívásokkal kellett szembenéznük, mint a keleten letelepedő honfitársaiknak. A legnagyobb kihívásokat, amelyeket a nyugati bevándorlóknak el kellett viselniük, a határhelyezettel kapcsolatos természetes nehézségek jelentették. E történészek kis variációkkal, de a Nyugaton élő írek komparatív sikerét mindnyájan annak tulajdonítják, hogy ezek az írek nem szembesültek azzal a vallási és etnikai intoleranciával, amely általános volt a keleti angolszász-protestáns közösségekben. Épp ellenkezőleg: a Nyugatot New Mexico, Texas és California államban felfedező írek az ott élő katolikus népességben olyan közösségekkel találkoztak, amelyek inkább befogadták, mintsem elutasították őket. Azok, akik Coloradóba és Montanába mentek, a bányákban bőséges munkát találtak, és egy olyan környezetet, ahol az embert első-sorban aszerint ítélték meg, mennyi követ rak a csillébe.³⁷ Ezek a feltételek nemcsak egyszerűen a nagyobb anyagi sikert alapozták meg, hanem, amint az adatok igazolják, nagyobb irodalmi termelékenységet, egy olyan irodalmat, amelyet különösen az ír perspektíva rögzítésének – atipikus mértékű optimizmussal teli – akarata jellemez.³⁸

Indokolt ugyan mindezen adatok részletesebb elemzése, de már ez az előzetes vizsgálódás is egy szabályos és megjósolható mintázatot tárt fel az ír-amerikai szerzők címadási gyakorlatában. Ez a mintázat a szerzőknek az Egyesült Államokba történő ír bevándorlás általános történetében elfoglalt kronológiai és földrajzi helyével függ össze. Az ilyen típusú elemzés alkalmas komparatív megközelítésre; más etnikumú bevándorlók által adott címek gyümölcsöző kontextust kínálnak. A zsidó-amerikai irodalomban hasonló mintázatot találnánk? Biztos, a két csoport között számos hasonlóság van irodalmukat és az Egyesült Államokba történő bevándorlással kapcsolatos tapasztalataikat tekintve. Az érdeklődő kutatónak mindössze azonosítania és kódolnia kell a zsidó-amerikai hagyomány bibliográfiai adatait azért, hogy lehetővé váljon egy ilyen elemzés.

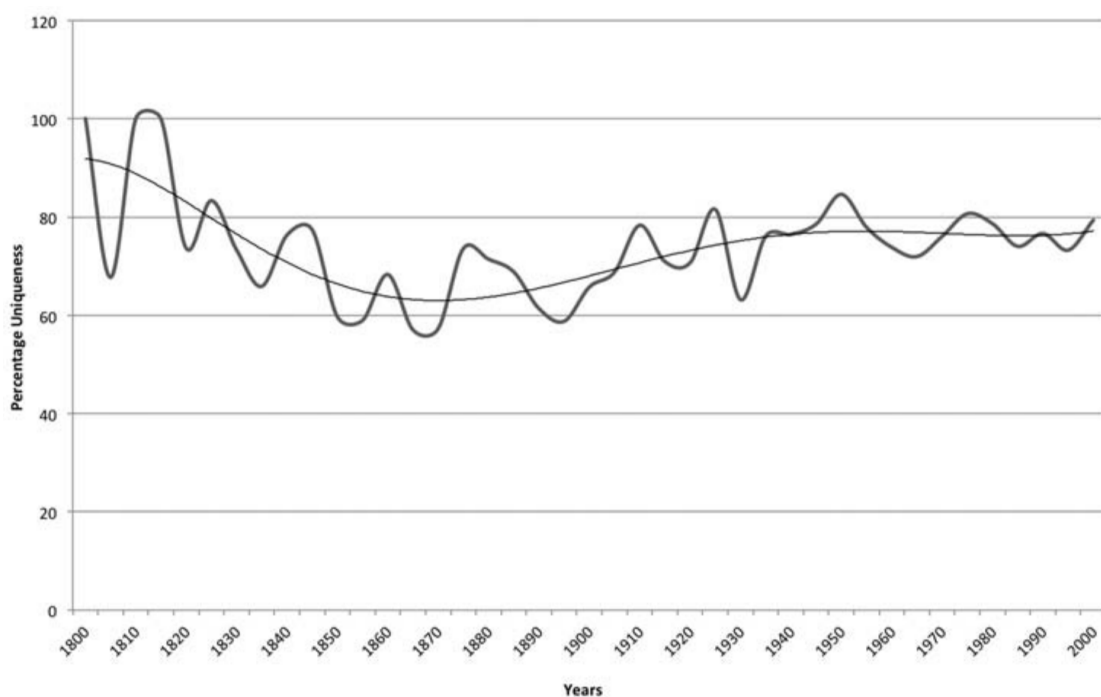
Sőt, párhuzamos bibliográfiák felállításával és összevetésével elvégzett összehasonlító elemzések változatos formáinak a segítségével továbbléphetünk a címekben szereplő szavak pusztá megszámlolásánál, és elemezhetjük a szavak aktuális használatát az érintett címeken belül. A 9-es ábra például grafikus reprezentációja a lexikai gazdagságnak – ez az egyik lehetőség a sok közül, mellyel a nyelvi változatosságot mérni lehet. A lexikai gazdagság, más néven a típus-jel-arány (*Type Token Ratio*,

³⁶ J. P. Walsh, *The San Francisco Irish, 1850–1976* (San Francisco: Irish Literary and Historical Society, 1978). D. M. Emmons, *The Butte Irish: Class and Ethnicity in an American Mining Town, 1875–1925*. Statue of Liberty-Ellis Island Centennial Series (Champaign: University of Illinois Press, 1989). D. M. Emmons, *Beyond the American Pale: The Irish in the West, 1845–1910* (Norman: University of Oklahoma Press, 2010). Blessing, „West among Strangers: Irish Migration to California, 1850 to 1880” (1977). P. J. Dowling, *California, the Irish Dream* (San Francisco: Golden Gate Publisher, 1988). P. J. Dowling, *Irish Californians: Historic, Benevolent, Romantic* (San Francisco: Scottwall Associates, 1998).

³⁷ A bevándorlást itt a férfiak dominálták.

³⁸ Az optimizmusról és a nyugati államokban élő írek általában vett sikereiről lásd Jockers, „In Search of Tir-Na-Nog: Irish and Irish-American Literature in the West” (1997). Jockers, „A Window Facing West: Charles Driscoll’s Kansas Irish,” 110–113. Jockers, „A Literature of Good Fortune,” 8–27. Jockers, „West of Éire,” 37–51.

TTR) a szavak egyedi előfordulási arányát fejezi ki egy adott szöveg összes szavához mérve. A lexikai gazdagságot két szerző szókincsének összevetéséhez használhatjuk. Herman Melville *Moby Dick*je 7.8-as *TTR*-rel rendelkezik, és egy nagyméretű szótárat, 16872 egyedi szóalakot (típust) tartalmaz. Jane Austen *Sense and Sensibility*jének 2-es a *TTR*-e és 6325 szóalakkal jóval kisebb a szótára. Austen minden egyes szót átlag 19-szer használ, míg Melville átlagban csak 13-szor. Ennélfogva a *Moby Dick* olvasva olyannak tűnik föl, mint amelyben az olvasó átlagban jóval több egyedi szót és kevesebb szóismétlést számolhat meg. Mindazonáltal ebben az egyenletben tekintetbe kell venni a regény hosszát is. Habár a Melville által használt szótár a *Moby Dick*ben jóval nagyobb, mint Austené a *Sense and Sensibility*ben, Melville regénye összesen 214889 szót tartalmaz, ami majd kétszerese a 120766 szót tartalmazó Austen-regény hosszúságának. A hosszúságbeli különbséget a két regényből véletlenszerűen választott 10000 szavas mintákkal kompenzálhatjuk, ha mindegyik mintának kiszámoljuk a lexikai gazdagságát, majd átlagoljuk az eredményeket. Austen esetében 100 véletlenszerűen kiválasztott 10000 szavas minta átlagos lexikai gazdagsága 0.19. A hasonló teszt Melville esetében 0.29-es átlagot hozott. Más szavakkal: a *Moby Dick* bármelyik 10000 szavas mintájában 2903 egyedi szóalak (típus) található, míg Austen mintái esetében csak 1935. Szótári értelemben Melville *Moby Dick*je jóval gazdagabb mű, mint a *Sense and Sensibility*.



9. ábra. Az ír-amerikai művek címének lexikai gazdagsága, kronológiai eloszlás

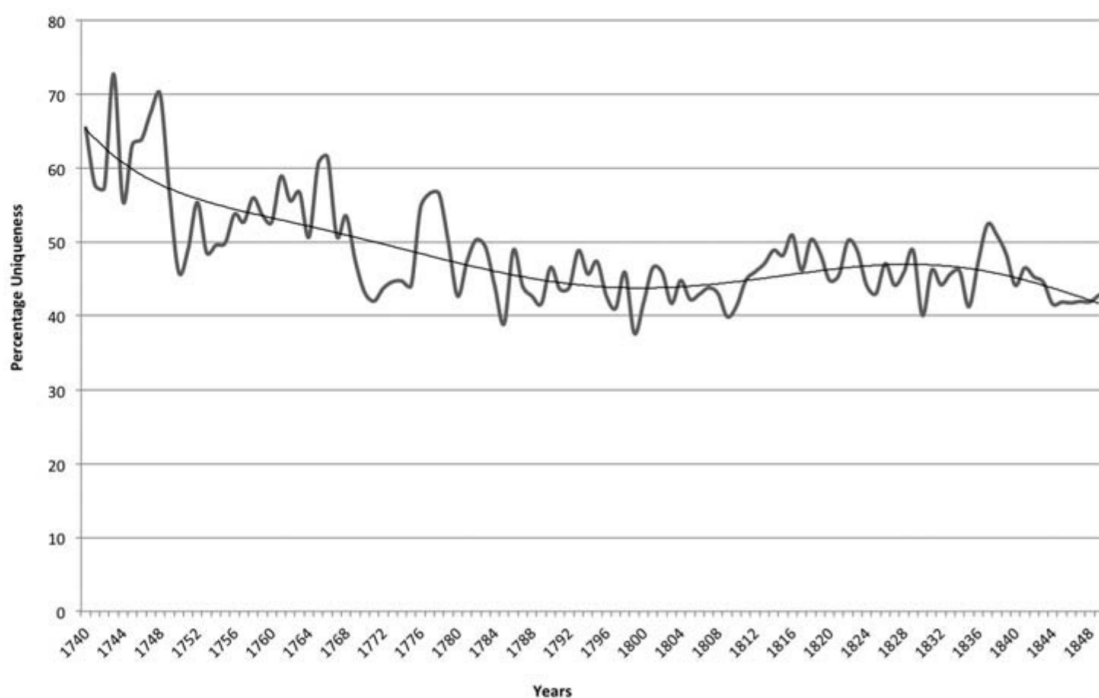
A művek címére alkalmazva a lexikai gazdagság mérését megvizsgálhatjuk, hogy egy adott időszak „címlexikona” milyen mértékben homogén. A címadatokat vizsgálva megbecsülhetjük egy adott időszakasz intervallumában a címek lexikai változatosságát. A 9-es ábra az ír-amerikai címek lexikai gazdagságát mutatja egy időtengelyen, és rögtön nyilvánvalóvá teszi a lexikai kiegyenlítődéssel történő haladást. A szabá-

lyosabb vonal a grafikonon a címek gazdagságának általános trendjét, tehát az átlagot mutatja, a másik, hullámosabb vonal láthatóvá teszi a kisebb változatosság felé vezető utat. Az átlagtól való eltérés az időben haladva csökken. A fluktuáció, az átlagtól való eltérés egyre kisebb lesz, és az 1930-as évekre ér el egy állandó szintet.³⁹ A korpusz korai éveiben a címek között nagyobb a változatosság; azaz a későbbi periódusokhoz képest magasabb az egyedi szavak százalékaránya. Praktikusan ez azt jelenti, hogy a lehetséges olvasó az 1820-as években a címeket böngészve nagyon különbözőnek találta őket (szótárakat és ennek következményeként feltételezhető tartalmukat tekintve), ellenben egy hozzánk időben közelebb megjelent szöveg olvasója nagyobb hasonlóságot talál e hipotetikus könyvespolc címei között. A kiegyenlítődés felé vezető általános útnál még érdekesebb – legalábbis az ír-amerikai irodalom iránt érdeklődők számára – a „gazdagság recessziója,” amely a nagy éhínséggel körülbelül együttjáró periódusban éri el mélypontját. Fanning, aki minden más tudósnál többet olvasott ebből az irodalomból, úgy jellemezte az ír-amerikai irodalom ezen korszakát, mint amelyből általában hiányzik a képzelődés. A korszak alacsony lexikai gazdagsága hasznos kvantitatív kiegészítés az ő kvalitatív megfigyelésének. Mégis tekintve, hogy viszonylag kisméretű a korpusz, a határozott következtetéseket nehéz statisztikai úton igazolni.

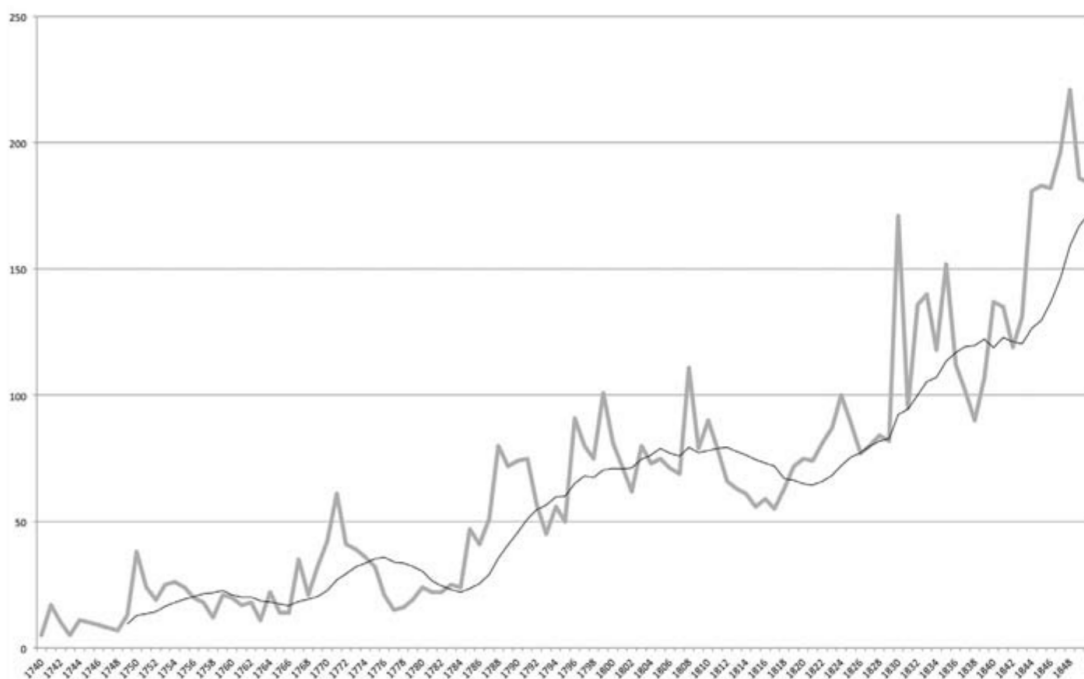
Egy nagyobb gyűjtemény esetén, mint amilyen Morettinek az 1740 és 1850 között publikált brit regényeket tartalmazó bibliográfiája, az időben változó lexikai gazdagság mérése jóval informatívabb. Moretti 110 évet felölelő bibliográfiája 7304 címet tartalmaz, azaz évi átlagban 66 címet. A 10-es ábra az időben változó lexikai gazdagság grafikus reprezentációja. 1810-ben, hogy egy önkényesen kiragadott példával éljek, 81 regény jelent meg. Mindössze egyetlen regény címe tartalmazta közülük a *Rackrent* szót: Maria Edgeworth *Castle Rackrent*-je. Másrészt a kastélyok elég népszerűek voltak 1800-ban. A *kastély* szó az 1800-ban megjelent 81 cím 10%-ában tűnik föl. A leggyakoribb töltelékszavak (*the, of, on, a, in, or, and, to, at, an, from*), valamint a műfajjelölők (*novel, tale, romance*) kiszűrése után valójában a *castle* a leggyakrabban előforduló szó. Ez nem teljesen meglepő, tekintve a gótikus regény műfajának 1790 és 1830 között betöltött jelentőségét, valamint azt a tényt, hogy a műfaj prototípusa, Walpole *Castle of Otranto*-ja az imitáció hosszú éveiben meghatározta a helyszínt, és berendezte a színpadot.

A 10-es ábra a típus-jel-arány évenkénti eloszlását mutatja. 1800-ban 275 egyedi szóalak (típus) volt és összesen 659 szó, 42%-os a gazdagság. Abban az évben átlagban 8 szó hosszúságúak, és átlagban 3.4 egyedi szót tartalmazott mindegyik. A szógazdagsági 10-es ábra a brit regények olvasói számára elérhető címek relatív homogenitását mutatja. A nagyobb képet nézve azt látjuk, ahogy haladunk előre az időben 1740 és 1840 között, a címek egyre kevésbé egyediek, a címekben használt szótár mérete csökken, és mindez abban az időben történik, amikor a publikált címek száma összességében drámaian megnő. A 11-es ábra a korpusz címeinek számát mutatja évenkénti eloszlásban, drámaian jelezve, milyen mértékben bővül a korpusz az időben előre haladva.

³⁹ Majdnem biztos, hogy a címek hossza is szerepet játszik. Ahogy haladunk a múltból a jelen felé a címek hossza az 1840 előtti durván 75 karakterről a 19. század második felében 50 karakterre esik, majd 1900 után 20–25-re.



10. ábra. A brit regénycímek lexikai gazdagságának kronológiai görbéje



11. ábra. A brit regénycímek évenkénti száma

Másképpen nézve, a szerzők által a címekben használt lexikon egyre homogénebb egy adott időszakban még annak ellenére is, hogy a piacon megjelenő könyvek és címek száma növekszik. Ha elfogadjuk Duchet és Moretti állítását, hogy a címek

fontos jelek vagy kódok egy tömegpiacon, akkor ez egy kijózanító trend, legalábbis azokra a szerzőkre nézve, akik jól akarnak szerepelni a könyvpiacra. Ahelyett, hogy a címek megkülönböztetnék a könyveket egymástól, csak a homogenizáció folyamatát látjuk, amelyben az egyre több könyv címe egyre inkább ugyanúgy néz ki. Moretti hasonló trendet állapított meg a címek hosszával kapcsolatban, de a speciális szótárral kapcsolatban nem vizsgálódott.⁴⁰

A *romance* szó például népszerű ebben a címkorpuszban. A töltelékszavak eltávolítása után valójában a tizenhatodik leggyakrabban előforduló szó a címekben. Ráadásul egy olyan szó, amely a népszerűség különböző korszakait érte meg: 1810 körül volt a csúcspontján, majd mintegy harminc évre kiment a divatból, mielőtt az 1840-es években ismét felemelkedett (lásd a 12-es ábrát).



12. ábra. A *romance* előfordulásának kronológiai görbéje a brit regénycímekben

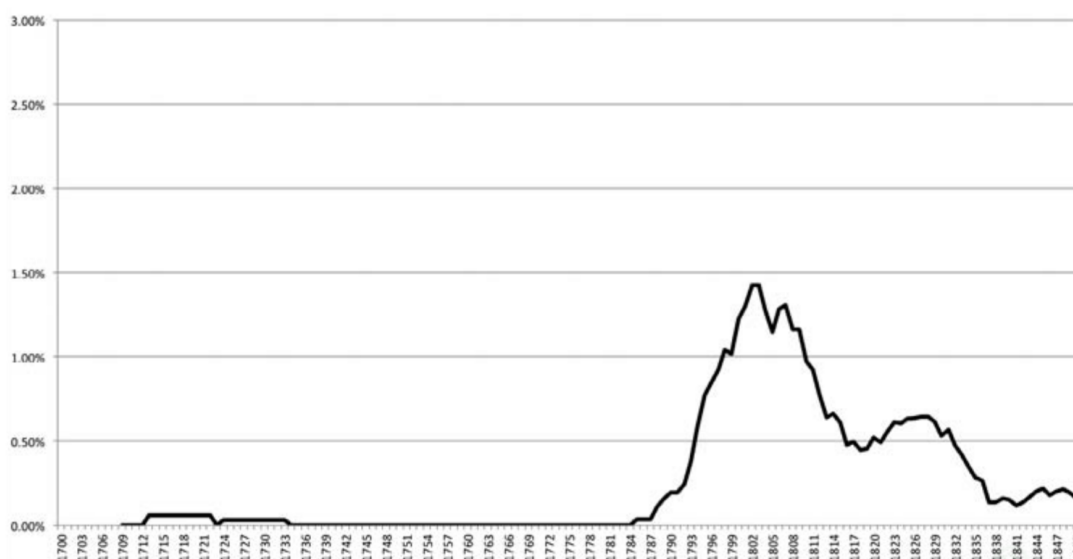
Van néhány magyarázat az itt látott trendre, de célom nem az, hogy a *romance* kismértékű növekedését megmagyarázzam, hanem, hogy ehelyett megmutassam azt a szélesebb kontextust, amelyben a *romance* szót tartalmazó címek feltűnnek, vagy még egyszerűbben, hogy megmutassam, hogyan néz ki a verseny. Azok az írók, akik az 1810-es években a *romance*-ot használták, kevésbé eredetiek, mint azok, akik az 1700-as évek közepén, végén alkalmazták. Azok a korai, kisebb kiugrások a *romance* időegyesítésén 1740 és 1760 körül az első adoptálók, akik eltérnek a fő trendtől; ők a kifejezés használatának úttörői, és bárki, aki meg akarja érteni a *romance* tömeges használatát a századfordulón, elemzését bizonyosan a trendet megelőzőkre vetett alaposabb pillantással kell kontextusba emelnie.

⁴⁰ Nyilvánvaló, hogy néhány szerző egyáltalán nem kívánt kitűnni a piacon. Épp ellenkezőleg, korábbi sikereket akartak tökévé alakítani, és a legnagyobb profitot eredményező címadási trendeket követni.

Ezzel összehasonlítva a *love* szó viszonylag kevésbé játszik a versenyben: egy meg-
lehetősen állandósággal használt címszó a korpuszban (13-as ábra), és sosem volt olyan
sikeres időszaka, mint amilyet a *romance*-nál megfigyelhettünk.



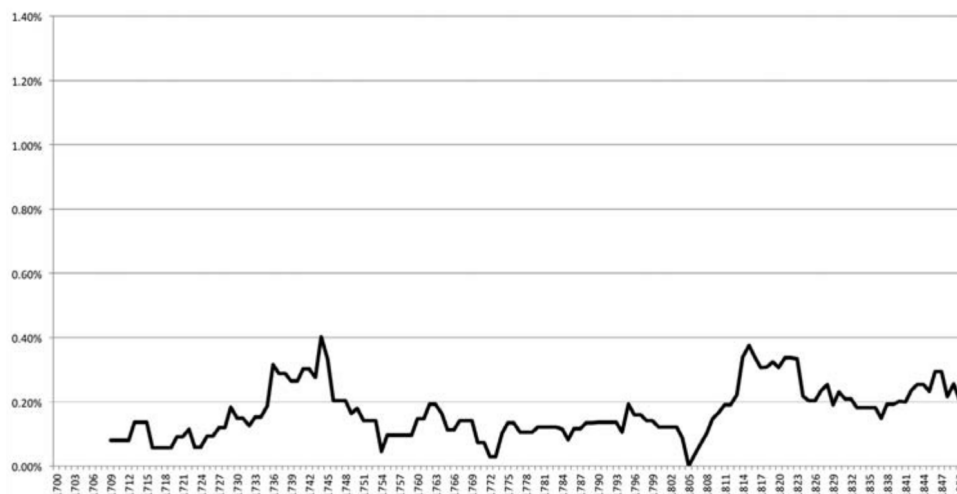
13. ábra. A *love* előfordulásainak kronológiai görbéje a brit regénycímekben



14. ábra. A *castle* előfordulásainak kronológiai görbéje a brit regénycímekben

Másrésről viszont a *castle* (14-es ábra) sokkal inkább úgy viselkedik, mint a *romance*, míg a *London* (15-ös ábra) a *love*-hoz hasonló. A *century* szó viselkedése (16-os ábra) nem meglepő: közvetlenül a századforduló előtt kezdődik az emelkedés, és közvetlenül utána, 1808 körül éri el a csúcst, amikor is hirtelen régi újdonsággá válik, túlhaladottá, klisévé. Az írekkel foglalkozó tudósok elég gondolkodnivalót találnak a

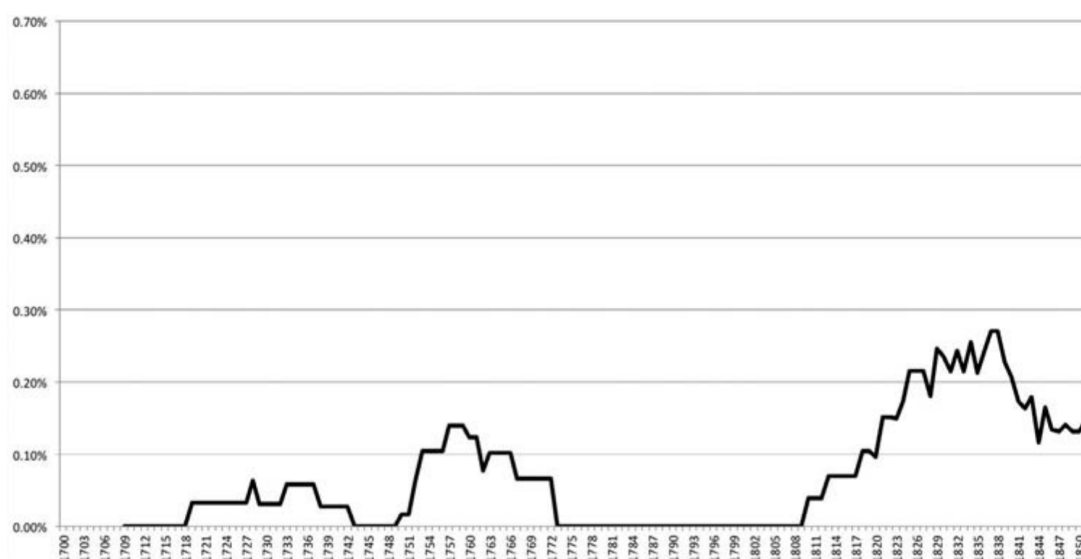
17-es ábrán, amely az *Irish* szó gyakoriságát ábrázolja ebben a címkorpuszban. Sosem volt különösebben népszerű kifejezés, csak 62-szer fordul elő, de amint látható, enyhe emelkedésnek indul a 18. század harmadik negyedében, és aztán egy erősebb, hirtelen az emelkedés a 19. század első negyedében. Mégis ami a leginkább elgondolkodtató, az a 18. század utolsó negyedének és a 19. század első évtizedének csöndje. Most beszéljünk egy elveszett nemzedékről! Bizonytalan, hogy ebben a peridusban miért volt nem népszerű téma a regénycímekben Írország és az írek, azonban az egyértelmű, hogy hirtelen eltűntek a könyvtárakból és a könyvesboltok polcairól.



15. ábra. A *London* előfordulásainak kronológiai görbéje a brit regénycímekben



16. ábra. A *century* előfordulásainak kronológiai görbéje a brit regénycímekben



17. ábra. Az *Irish* előfordulásainak kronológiai görbéje a brit regénycímekben

Az ebben a fejezetben tárgyalt adatok több kérdést vetettek fel, mint ahányat megválasztak. Ez a tény megnyugvással tölthet el mindenkit, aki amiatt aggódik, hogy az irodalom mérése megkondítja a lélekharangot a további tanulmányok felett. Mindazonáltal felbukkant legalább egy érdekes szál ebben a kezdeti elemzésben. Felvetette a külső hatások kérdését, valamint azt, hogy az olyan faktorok, mint a nem, nemzetiség, idő, földrajzi helyzet milyen mértékben határozzák meg a szerzők döntését arról, hogy miről írnak, és miként adnak címet munkájuknak. Feltűnő, hogy a nem, a földrajzi elhelyezkedés és az idő a regénytermelés és a regénycímek befolyásoló tényezői, valamit befolyásolják magukat a témákat is, amelyekről a szerzők saját elhatározásukból írnak. De a címek és a gazdag metaadatok vizsgálata csak ennyire ad lehetőséget. A következő fejezet ezeket a kérdéseket tárgyalja a nyelvi stílus elemzésén keresztül, mégpedig a 19. századi fikciós próza tizenkét népszerű műfajából 106 teljes szövegű regénynek a korpuszán.

Fordította: Labádi Gergely

Metadata

Chapter 5 of *Macroanalysis: Digital Methods and Literary History* demonstrates how computational approaches can be used to explore bibliographic and demographic metadata in order to reassess and postulate new claims about literary history. This chapter demonstrates how thought-provoking perspectives on literary periods and trends can be derived from a macroanalysis of book-level metadata – including book titles, author ethnicities, publication years, fictional settings, etc. Utilizing a database of metadata related to Irish-American literature, the author provides a reassessment of common narratives of Irish-American literary history and offers an alternative perspective on Charles Fanning's theory of a "lost

generation” of Irish-American writing. The metadata provide context for understanding Fanning’s reading of literary history and suggest that scholarly assumptions about Irish-American literary history might be based too much on a small sample of books by a homogenous group of authors. More broadly, the chapter argues that traditional literary scholars make a mistake in thinking that large-scale analyses are proposed as a replacement to close reading. On the contrary, a macroanalysis of metadata provides the necessary context for close readings and prompts new questions and new perspectives on literary history.

Keywords:

macroanalysis, metadata, distant reading, close reading, Irish-American literature

<MŰHELY>

Sennyey Pongrácz

St. Edward's University Library, Austin, Texas

pongacz@stedwards.edu

Viták és víziók a digitális bölcsészetről

A cikk áttekinti az Amerikai Egyesült Államokban lezajlott digitális bölcsészet körüli legújabb vitákat, és összefoglalja azok főbb érveit. Az írás számos kérdést kiemel e vitából, például hogy milyen kritériumokat kell egy projektnek teljesítenie ahhoz, hogy azt digitális bölcsészettnek lehessen minősíteni, valamint hogy a digitális bölcsészet tudományának vagy inkább módszertannak tekinthető. Új infrastruktúrát kell kifejleszteni ahhoz, hogy a kutatók megkaphassák a szükséges támogatottságot digitális bölcsészeti projektek megvalósíthatóságához. Az írás végül azt taglalja, hogy a jövőre vonatkoztatva az e tárgykörbe tartozó diskurzusoknak milyen következményei lehetnek a bölcsészettudományokra.

Kulcsszavak:

digitális bölcsészet, digitális infrastruktúra, digitális módszerek



Amióta a digitális bölcsészet fogalma megjelent, viták vannak körülötte. Ezek a viták hol a kifejezés szabatos értelmezéséről szólnak, hol arról az alkalmazott eszközkészletről, amelyeket e terület kutatói alkalmaznak, hol pedig az ebben a kontextusban született új tudományos eredményekről folyik a diskurzus. Ezeknek a kérdéseknek a tárgyalása Magyarországon sem ismeretlen, konferenciákon és szakcikkekben is találkozni velük. Az látszik belőlük, hogy e kérdések egyre kifinomultabban és árnyaltabban újra és újra felmerülnek. Az elmúlt években ezek a viták felerősödtek az Egyesült Államokban, és különböző megközelítések jelentek meg úgy a tudományos folyóiratokban, mint a témához kapcsolódó hetilapokban is.

A vitákat érdemes követnünk Magyarországon is, mert nagyon tanulságosak lehetnek a hazai szakemberek számára akkor is, ha még az amerikai kollégák sem találták meg a bölcsék követ. A diskurzusokat – az áttekinthetőség kedvéért – *öt alapvető témára* lehet bontani, ha ezzel némiképp le is egyszerűsítjük a valójában igen komplex megszólalásokat.

Eszközhasználat?

Az első téma, amelyről vita bontakozott ki, az alkalmazott eszközökről szól.¹ Az a kérdés, hogy milyen eszközök használatával válik egy tudományos projekt digitális

¹ Timothy Brennan, „The Digital-Humanities Bust,” *Chronicle of Higher Education*, 2017. okt. 15., <http://www.chronicle.com/article/The-Digital-Humanities-Bust/241424>.

bölcsészeti. Ha számítógépe(ke)t alkalmaz a kutató a munkában, az már elegendő hozzá? Vagy szükséges valamiféle speciális algoritmust is futtatnia? S egy ilyen célszoftver alkalmazása elegendő kritérium? Felmerül az a kérdés is, hogyha például a kutató létrehoz egy digitalizált gyűjteményt (amelynek az eredetije analóg formátumban már létezett), vagy ha egy korpuszt például XML-formátumba rögzít, akkor az ilyen jellegű projektek digitális bölcsészetnek számítanak? Más szóval, kell-e különbséget tenni a digitális eszközök bölcsészeti használata és magának a digitális bölcsészetnek a fogalma között? Lehet-e pontos határt szabni, kritériumokat kijelölni, hogy jól pozicionált distanciákat tudjunk tenni e két metódus között? Az mindenesetre látszik, hogy eddig nem sikerült a pontos határvonalat meghúzni. Persze az is kérdés, hogy fontos-e pontosan kijelölnünk.

Módszertan?

A második vitatéma az adat (vagy forrás) feldolgozási módszerei körül forog.² Miféle új módszereket tud a digitális bölcsész alkalmazni, és azzal milyen új eredményeket tud felmutatni? Mivel a digitális bölcsész számos példát hoz az új kvantitatív módszerek alkalmazásával kapcsolatban (amivel mind nagyobb forrásanyagot tud feldolgozni), ez a vita alapvetően skálákról, skálázhatóságról szól. Az új módszerek (és eszközök) előnye gyakran abban ragadható meg, hogy eddig leküzdhetetlen méretűnek számító szöveg- és adathalmazokat kezelhetővé tud tenni. Gépek és algoritmusok nélkül ezeket lehetetlen lenne/lett volna belátható időn belül feldolgozni. Az is realitássá vált, hogy a digitális bölcsész – az eddigi esetleges benyomások és vélelmek helyett – korábban elérhetetlen pontossággal és matematikai precizitással rajzolja fel az eredményeket. Ebben az összefüggésben az további vita tárgya lehet, hogy a tudós ilyenkor csak azt tudja az algoritmusával feldolgoztatni, amit ő maga illesztett be az algoritmus programozásakor – hiszen ma még (!) a kutatónak és nem a számítógépnek a kreativitása és intelligenciája, illetve képessége áll az eredmény mögött. Emiatt kerülnek elő új összefüggések, rajzolódnak ki új mintázatok, válnak láthatóvá új kapcsolódások. Napjainkban még nem tartunk ott, hogy a mesterséges intelligencia segítségével elemezni lehessen ennyire összetett problémákat. Ezek az absztrakt lépések továbbra is a kutató képességeitől (és lehetőségeitől) függenek. Az is igaz ugyanakkor, hogy attól még, hogy a kutató megalkot új fogalmakat, feltesz új kérdéseket, abból még nem feltétlenül következik, hogy be is tudja illeszteni azokat az algoritmusába. Valójában két olyan határról van itt szó, amelyeket egyszerre kellene átlépni: egyrészt a nagy mennyiségű adatok új összefüggéseinek a feltárását, másrészt az informatikai megvalósítás mikéntjét.

Gondolkodásmód?

Az előbbi viták – gyakran hallgatólagosan – arról is szólnak, hogy a digitális bölcsész egy újféle gondolkodást hoz előtérbe vagy akár meg is teremt, illetve máshonnan nézve

² Saulo Cunha de Serpa Brandão and Wander Nunes Frota, „On the Path to a Methodology for the Critique of Digital Literature,” *Digital Scholarship in the Humanities* 32, 2. sz. (2017): 225–233, <https://doi.org/10.1093/llc/fqw016>.

megköveteli azt. Olyat, amely a természettudományok pontosságát és precizitását vezeti be a humán tudományágakba. Ez az úgynevezett „algoritmikus gondolkodás” (computational thinking).³ Ez több szálon visszavezet a „strukturalista iskolák” céljaihoz, amelyek még a 20. század elején, az informatika kibontakozása előtt jelentek meg. Sőt, már korábban is léteztek ilyen tendenciák, láthatóak voltak ilyenféle szemléleti megközelítések, gondoljunk például a híres francia *Annales* folyóiratra, de akár már a pozitivizmus aranykorában is vannak erre a gondolkodásmódra példák, ilyenek Szinyei József (1857–1943) hatalmas adatmennyiségeket megmozgató módszerei (nem véletlen emlegették „gőzhangya”-ként kortársai). A megközelítésben voltak hasonlóságok, de persze akkor még kézzel dolgozták fel a forrásanyagot, gyakran hosszú éveket szánva a munkára. Az ilyen tudós – kortársai többségétől egészen eltérő módon – alapvetően máshogyan közelítette meg az anyagát, hiszen számokkal, statisztikai képletekben, eloszlásokban, táblázatokban gondolkodott. Az eredményeket gyakran nem folyamatos narratívákban fogalmazta meg, hanem adatok, grafikonok, táblázatok sorát tárta az olvasó elé, melyek összességükben adtak egészen új megközelítést az adott tudományon belül (tipikus korai és nagyszerű példa erre Fernand Braudel klasszikusa *A Földközi-tenger és a mediterrán világ II. Fülöp korában* három kötete).⁴

Külön tudomány?

Olykor a vitában egy harmadik dilemma is felmerül: vajon a digitális bölcsészet maga egy külön tudományág?⁵ Ha már a bölcsész tudósok számítógépeket használnak, algoritmusokkal dolgozzák fel az anyagukat, és ebből fakadóan *de facto* máshogyan is gondolkodnak, akkor az már valószínűleg egy külön tudományágot is jelent – vagy mégsem? Itt persze megkerülhetetlen az a tény, hogy attól még, hogy a digitális bölcsészek máshogy előkészített és új módon megjelenített alapanyagokat dolgoznak fel, és más aspektusból közelítenek hozzájuk – legyenek azok bár történelmieket, irodalmiak vagy éppen művészeti stb. –, még akkor is a bölcsészettudományokon belül maradnak és dolgoznak. Érdemes itt a tudománytörténet néhány analógiájára gondolni. Mi történik, ha egy tudományág bevett módszertanában egy egészen új terület bukkan fel? Például a biokémia, kísérleti fizika stb. sorra fontos szakterületekké váltak, de azért az alaptudomány (biológia, kémia, fizika) részei is maradtak, csak egyfajta szakosodást jelentettek. Valójában nem ritka, hogy egy tudományágban – éppen az új metodológiák miatt – újabb tudományterület alakul ki, de az ellenkezőjére

³ Adam Kirsch, „Technology is Taking over English Departments,” *New Republic*, 2014. máj. 2., <https://newrepublic.com/article/117428/limits-digital-humanities-adam-kirsch>. David M. Berry and Anders Fagerjord, *Digital Humanities. Knowledge and Critique in a Digital Age* (Malden, MA: Polity, 2017), 40–60.

⁴ Kokas Károly, „Digitális bölcsészet 2016. A bölcsészek és az informatikai megközelítés: régen és most,” *MONOKgraphia: tanulmányok Monok István 60. születésnapjára*, szerk. Nyerges Judit, Verók Attila, Zvara Edina (Budapest: Kossuth Kiadó, 2016), 405–412, <http://publicatio.bibl.u-szeged.hu/10296/>. E tanulmány a 2015-ben megrendezett „Digitális bölcsészet Szegeden” című minikonferencia előadásaira és tanulságaira is támaszkodik. Hozzáférés: 2017.09.15, <http://digibolcsesz.ek.szte.hu/>.

⁵ Alan Liu, „Is Digital Humanities a Field? An Answer from the Point of View of Language,” *Journal of Siberian Federal University. Humanities and Social Sciences* 7 (2016): 1546–1552, <https://doi.org/10.17516/1997-1370-2016-9-7-1546-1552>.

is van példa: egy új metodológia megváltoztathat egy tudományágat akár úgy is, hogy nem válik külön tőle, mint például a lézertechnika és a sebészet összekapcsolódása. Összegezve azt állíthatjuk, hogy ma még a digitális bölcsészet pontos helye a tudományterületeken belül, illetve hatása az egyes bölcsészeti tudományágakra nem pontosan tisztázott, és ez a helyzet még valószínűleg egy ideig így is marad. Végül érdemes megjegyezni azt is, hogy vannak olyanok, akik már eleve úgy fogalmazznak: a digitális bölcsészet pontos mibenléte vitatott terület marad a jövőben is.⁶

Divat?

A legutóbbi vitákban olyan megközelítések is felmerültek, hogy a digitális bölcsészet inkább csak egy *divattermék*, amelyet az egyetemi vezetők, tudományszervezők és kutatóintézetek használnak kedvezményezett pozíciók elérésére, a tradicionálisan dolgozó kollégáik rovására.⁷ Ezáltal, főleg az Egyesült Államokban, többen úgy érvelhetnek, hogy ők a „korszerűbb” digitális bölcsészetet művelik, és a fentiekben felvázolt homályos határokat is figyelembe véve, talán nem is tartoznak a szorosabban vett digitális bölcsészet égisze alá. Miután a sikeres (vagy szerencsés) bölcsész így hozzáfér olyan kutatási ösztöndíjakhoz és támogatásokhoz, amelyeket esetleg a hagyományos módszerekkel dolgozó kutatótársai nemigen kapnak meg, ez önmagában vitát eszkalál a „régiek és újak” közt a források elérhetősége tekintetében is.⁸ Nem kevés iróniával közelítik meg ezt a kérdést a konzervatívabb tudósok, akik a digitális bölcsészetben nem látnak mást, mint csupán időleges fellángolást, furcsa hóbortot. Szerintük a digitális bölcsészet nem különbözik a posztmodern vagy a multikulturális, esetleg a posztgyarmati témák és hasonló felbukkanásától, és olyanok lesznek, mint más „tudományos divatok”, amelyek aztán rendre rövid életűeknek bizonyultak.⁹ Ezek a megnyilatkozások igen lenézők, és valójában nem vitatják, de el sem ismerik a digitális bölcsészet alapvetéseit és annak összetett természetét.

Együtt

Olyan véleményekkel is találkozni azonban, amelyek a digitális bölcsészetet éppen az összefüggésekben találják meg. Bármelyiket nézzük is a fentiek közül, a digitális bölcsészet egy tudományterület és egy módszertan házasságából születik meg, amelyben talán minden eddiginél fontosabb szerepet kap a különböző kutatók közötti

⁶ Berry and Fagerjord, *Digital Humanities*, 148–149.

⁷ Carl Straumsheim, „Digital Humanities as a »Corporatist Restructuring«,” *Inside Higher Ed*, 2016. máj. 6., <https://www.insidehighered.com/news/2016/05/06/article-criticizes-impact-digital-humanities-colleges>.

⁸ Daniel Allington, Sarah Bouillette, David Golumbia, „Neoliberal Tools (and Archives): A Political History of Digital Humanities,” *Los Angeles Review of Books*, 2016. máj. 1., <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>; Alan Jacobs, „Critiquing the Critique of Digital Humanities,” *The New Atlantis*, 2016. máj. 2., <http://text-patterns.thenewatlantis.com/2016/05/critiquing-critique-of-digital.html>.

⁹ Stanley Fish, „The Old Order Changeth,” *New York Times*, 2011. dec. 26., <https://opinionator.blogs.nytimes.com/2011/12/26/the-old-order-changeth/>.

együttműködés. Ilyen módon – a bölcsészeti entitások közül – kiemelkedően függ a tudományos hálózattól is. Ezek a komponensek együttesen új eredményeket produkálnak, amelyeknek lényeges attribútuma, hogy új formában jelennek meg. Az egész „építmény” fontos jellemzője, hogy technikailag és strukturálisan egy új infrastruktúra köti össze részeit és szereplőit.¹⁰

A digitális bölcsészet kritériumai

Ezek a diskurzusok a lényegi problémákra fókuszálnak a digitális bölcsészet területén, olyanokra, amelyekkel fontos nyíltan szembenézni, és érdemes világosan elősorolni az érveket és ellenérveket is. E viták kulcsfontosságúak lehetnek a bölcsészettel foglalkozó szakmák megújulásához és a fejlődés fenntarthatóságához, még akkor is, ha láthatóan a kritikai megszólalások gyakran összekeverik a lehetséges megközelítéseket, és ezáltal bonyolultabbnak és nehezebbnek tűnik a jó válaszok megtalálása. Ezek a viták nem fognak egyhamar megszűnni: számítani lehet arra a forgatókönyvre is, hogy hosszú távon velünk maradnak. Fontos hatásuk, hogy újfajta gondolkodást ösztönöznek, továbbá a fejlődő metodológiákkal kapcsolatos állásfoglalást és a felhasznált eszközök folyamatos javítását és tökéletesítését is eredményezik. Ez a folyamat nyilván az érintett szakterületek mindegyikére pozitív hatást gyakorolhat, érdekesebb és eredményesebb új utakra is vezethet.

Ahhoz, hogy ezeknek a vitáknak hozadéka legyen, bizonyos feltételeknek teljesülnie kell. Elsősorban meg kellene határozni, hogy mi az a határ vagy mérőföldkő, amelynek átlépése után a tudós egyértelműen a „digitális bölcsészet” területére kerül. Másodsorban fontos lenne megvitatni, hogy a digitális bölcsészetet „csak” egy új módszernek vagy „már” egy új tudományágnak tekintjük. Ez utóbbi – mint láttuk – „politikai döntés” egyben, amelynek szakmapolitikai és támogatási következményeivel is számolni kell. Harmadik feltételként ki kellene fejleszteni egy olyan infrastruktúrát, amelynek segítségével a digitális bölcsészet művelésének területei és az alkalmazott technikák használata nem hárulna egyedül a kutatókra, hanem platformszerű általános segítséget és integrált eszközöket kapnának. Ezáltal a terület és a módszer hasznosíthatósága többek számára elérhetővé válik, és a projektek hosszú távon is továbbfejleszthetők maradnak. Végül fontos lenne kicsit a jövőbe pillantva átgondolni, hogy az előbbi válaszoknak mik lesznek, lehetnek a következményei a bölcsészeti tudományokban.

Kritériumként a következőket vethetjük fel, akár vitaindítónak is:

- A létező forrásanyag feldolgozható legyen digitálisan, hogy új szemszögből lehessen az anyaghoz hozzáférni, elemezni és vizualizálni. A bölcsészeti eredeti forrás – jelentős részben – írott szövegekből fog állni. Ezt a digitális bölcsész fel fogja dolgozni olyan formában, hogy ezt az anyagot és a végrehajtott műveletek eredményeit a folyamat végén grafikailag is meg lehessen „mutatni”, nem

¹⁰ Matthew Kirschenbaum, „What Is Digital Humanities and What’s It Doing in English Departments?” in *Debates in the Digital Humanities*, eds. Matthew K. Gold and Lauren F. Klein, 3–11 (Minneapolis, Minn.: University of Minnesota Press, 2012), 4, <https://doi.org/10.5749/minnesota/9780816677948.001.0001>.

csak szöveggént. A grafikai megjelenítés alatt gondolhatunk földrajzi térképre, gráfokra, vektorokra, grafikonokra stb.

- A forrásanyag olyan mennyiségi paraméterekkel rendelkezik, hogy a kutatónak számítástechnikai eszközöket kell alkalmaznia ahhoz, hogy a korrekt és időben belátható feldolgozásra egyáltalán remény legyen. Ezért új eszközöket és új metodológiát kell alkalmaznia. Ellenérvként felhozható, hogyha csekély méretű forrásanyagot használ a kutató, akkor a hagyományos eszközök és módszerek is megfelelnek az anyag feldolgozásához, viszont amennyiben a forrásanyag mérete átlép egy bizonyos küszöböt, a hagyományos eszközök már nem lesznek elegendők. Példának okáért az Árpád-kori fennmaradt forrásanyagot hagyományosan még fel lehet dolgozni, de már az Anjou-kori anyag túllépi ezt a határt, és itt már új eszközökre van szükség a teljes értékű és korszerű feldolgozáshoz. Megjegyzendő, hogy Magyarországtól nyugatabbra ez a határ korábban jelenik meg, többek közt az írásbeliség fejlettebb volta és a nagyobb méretű „forrásmegmaradás” okán. Mindez nem független attól, hogy a középkortudomány komoly forrás- és információs „hézagokkal” küzd régóta Európa-szerte.
- A forrásanyag megfelelően strukturált és szegmentált legyen ahhoz, hogy lehetőségessé váljék a gépesített elemzése. Vagyis az anyagot elő kell készíteni ahhoz, hogy felosztása, szervezése s főképpen metaadattal való ellátottsága lehetővé tegye az algoritmikus feldolgozást. Nyilván az eredeti forrásokra nagyon komoly előmunkálatok várnak, mielőtt azokon algoritmusokat lehet futtatni. Ez a nehézség nem alábecsülendő, és ezen a ponton szorul a digitális bölcsész komoly informatikai tudásra és segítségre. Persze ideális esetben magának az anyagnak az előkészítését is lehet számítógépes eszközökkel elősegíteni. Mondhatnánk azt is, hogy ez a lépés az, amely megkülönbözteti a *big data* típusú előfeldolgozott anyagot a nyers adathalomtól.
- A forrásanyag feldolgozásához szükséges folyamatokat automatizálni lehessen, és új minőségben, új szemszögből tudja a kutató vizsgálni a szövegeket. Miután a forrásanyagot algoritmikusan fel lehet dolgozni, maguk az algoritmusok is tudományos fontosságot nyernek. Azoknak a fejlesztése, minősége és használata is tudományos eredménnyé, továbbá szakmai viták témájává válik. Újabb szempontként felmerül, hogy a vizsgált anyagot más algoritmusokkal később ismét fel lehessen dolgozni, akár más kutatók közreműködésével is. (Természetesen ez a folyamat az *open data* problematikáját is felveti.)

Azt mondhatjuk, ha ezen kritériumok közül *legalább kettő* teljesül, akkor talán lehet arról beszélni, hogy az adott projekthez hozzárendelhető a *digitális bölcsészet* címke.

Együtt és külön

A feltett kérdésekhez fontos még figyelembe venni azt is, hogy a digitális bölcsészet módszere több hagyományos tudományágban alkalmazható. Ennek a ténynek a figyelmen kívül hagyása azzal a veszéllyel fenyeget, hogy egy interdiszciplináris irányt indítunk el, amely akarva akaratlanul oda vezethetne, hogy a jövőbeni erőforrások

szétforgácsolódnak. Márpedig amit a történész, a nyelvész, az irodalmár stb. el akar érni a fent említett kritériumok alkalmazásával, az igen eltérő is lehet. A metodológiákban lehetnek (és valószínűleg lesznek is) hasonlóságok (amint a hagyományos metodológiákban is megfigyelhetők ezek), de a célok, a korlátok, a folyamat és a problematika más és más lesz. Sőt, a háttértudás, amely nélkül nem lehet tudományos haladás, ugyanannyira eltérő a digitális anyagok kezelésében, mint az analóg világban. Ezért azt érdemes átgondolni, hogy nem járna-e jobban mindenki, ha nem külön tudományágként közelítenénk meg a lehetséges újításokat, amelyeket a digitális bölcsészet nyújt. Esetleg jobb volna az erőfeszítéseinket két kihívásra fókuszálni: arra, hogy a hagyományosan dolgozó szakmák jobban elfogadják a módszertani újításokat, amelyeket a digitális világ nyújt, illetve hogy teremtsük meg a feltételeit és a lehetőségét annak, hogy a közösen alkalmazható eszközöket, kapacitásokat, algoritmusokat és stratégiákat meg tudják osztani egymás közt a tudósok.

Együttműködés

Az is fontos tapasztalat e területen, hogy a legtöbb digitális bölcsész ab ovo abból indul ki, saját magának kell megszereznie az alkalmazandó eszközök használatáról való tudást, nemritkán az algoritmusok megtervezését (néha megírását) is, emellett nem elhanyagolható az adott forrásanyag ismerete és feldolgozási terve sem. Ezek a kritériumok összességében csak igen kevés és nagyon komplex tudású szakembert feltételeznek. Sajnos emiatt az ilyen jellegű tervek vágyak maradnak. Félresikerült torzókká válhatnak olyan projektek is, amelyek azonban kellő feltételek mellett sikeresek lehetnének. Minthogy a technikai fejlődés megállíthatatlan, és a digitális bölcsészet ezzel a haladással igen szoros függő viszonyban van, nem lehet elvárni, hogy a bölcsész ne csak a saját szakmájával és annak forrásaival tartson lépést, hanem annak teljes informatikai vertikumával is. Talán ennek az igen erős kihívásnak mégis vannak megoldásai. Kezdjük azzal, hogy az egyetemen a bölcsész nem magányosan létezik, például az adott egyetemen belül vannak más tanszékek, karok és intézetek, amelyek már rendelkeznek ismeretekkel ezen a területen. Lehet, hogy olyan nyersanyagra várnak vagy kooperáló partnerre, aki tartalommal tudja kiszolgálni elméleti konstrukciójukat. Ilyen partner lehet a könyvtár és/vagy az informatikai tanszék vagy intézet.¹¹ Nagyon nagy szükség volna a különféle kooperációkra, talán együttműködési szerződéseket kellene kidolgozni, amelyek a felek számára megfelelő előnnyel bírnak. Egy ilyen „védernyő” alatt a közös projektek hosszú távú fenntarthatóságának esélyei is növekednének.

A jövő víziója

A jövőbe tekintve látszik, hogy ezen a területen a tudós még kevésbé lehet magányos, mint másutt. A digitális bölcsészethez szükséges megközelítések többet nyernének, mint veszítenének, ha a szükséges részzakterületek kutatói néha bepillantathatnának egymás munkájába. Például az e-science-es és a digitális bölcsész sokat tanulhatna egymástól. A megfelelő algoritmusok megtervezése és a programok megírása, a *big*

¹¹ Berry and Fagerjord, *Digital Humanities*, 80–81.

data szemlélet és a szövegkataszterek feldolgozása közös kihívásokat rejt, és hasonló megoldások után kiált.

Az újonnan létrejövő digitális bölcsészetnek (tartsuk azt akárminek is) az elnevezése, mint „e-”, illetve „digitális”, nem túl szerencsések. Elvégre a technika haladása mindig új kapukat nyit előttünk, és lehetetlen azt gondolni, hogy a tudományos munkát lehetséges lenne az aktuális legfejlettebb technika bevonása nélkül folytatni. Magától értetődő, hogy itt, a tudomány vizsgálódási területein kerülnek először bevonásra ezek az új technikák és technológiák. A valódi és fontos kérdés az, hogy a fejlődő technikát hogyan lehet az egyes tudományokban *legjobban és legeredményesebben* felhasználni. A lehetőségek, az előnyök és az új módszerek állandó keresését és kidolgozását természetes fejleménynek kellene tekinteni. Ugyanakkor el kell fogadnunk azt is, hogy az újjal szemben erős a szkepticizmus, a kritikai él, és az ellenállás magasabb szintje természetesnek mondható, mivel a bölcsészeti tudományok éppen a technológiákkal szemben előítéletesek és egyébként is meglehetősen konzervatív szemléletűek. Ez önmagában nem újdonság, nem is a mi korunk szülötte, hiszen távoli és közeli tudós újítások mind-mind ezzel a kérdéssel küszködtek Galileitől Einsteinig.

Az analóg és digitális módszertanú tudományok elválasztása sokszor inkább a tudományos iskolák és műhelyek szakpolitikai súrlódásaira vezethető vissza, hiszen a kutatói előmenetel, a projektfinanszírozás elnyerése, a pályázatok, a presztízs és hírnév a módszertanok megkülönböztetéséhez vezet, és mindez súlyos érdekellentéteket generál. Ezért is fontos tudomásul venni, hogy a digitális bölcsészet egyfajta függő állapotban van, ha úgy tetszik, állandóan kiszolgáltatott az informatikai feltételeknek és szaktudásnak. Vagyis arra kell számítani, azzal kell együtt élnünk, hogy a digitális bölcsészet az állandó megújulás és változás mellett állandó partnerkeresésben leledzik majd, mert a legújabb technikákat és a legjobb tudást – ami megvalósítja elképzeléseit, amivel lépést tud tartani – csak így kaphatja meg. Csak reménykedni lehet abban, hogy a lehetséges partnerek a másik oldalról hasonlóan gondolják mindezt, s nekik is szükségük van arra a tudásra és főként anyagra, amelyet a digitális bölcsész tud előállítani.

Az is látszódik ma már, hogy egyre hatékonyabbak az új eszközök és metodológiák, amelyeknek köszönhetően új eredmények jönnek létre. Olyanok, amelyek új kérdéseket tesznek föl, s néha régieket válaszolnak meg. Ez ma már igazolja e törekvések létjogosultságát. S az is meglehet, hogy előbb-utóbb nem kap majd különösebb megkülönböztetést az a projekt, amelynek eredményei részben vagy egészben a ma digitális bölcsészetnek nevezett szemlélettel és módszerrel keletkeztek. Innen nézve a tudományos eredmények lesznek a megkülönböztető és fontos tényezők, nem a felhasznált technika és módszer. Akkorára sikeresnek fogják elkönyvelni a digitális bölcsészet erőfeszítéseit, azok helyükre kerülnek, és a manapság zajló és élénk viták csupán felejtető háttérzajnak bizonyulnak.

Discussions and visions of the digital humanities

This article reviews the latest controversies surrounding digital humanities in the United States and summarizes their main arguments. It discusses a set of questions that these controversies raise, starting with the lack of definition as to what criteria need to be met to qualify a project as a digital humanities project. It discusses whether the digital humanities is a methodology or a new branch of scholarship. It argues for the need to develop an infrastructure to support the pursuit of digital humanities projects. It concludes by identifying future implications of digital humanities research for scholarship in the humanities.

Keywords:

digital humanities, digital tools, digital methods

Horváth Iván

Eötvös Loránd Tudományegyetem, Régi Magyar Irodalom Tanszék

ivan@elte.hu

Digitális bölcsészet a virtuális nemzeti könyvtárban

Egy virtuális nemzeti könyvtár tetszőleges példányszámban szétsztható az emberek között. A sokszorosítás a hosszú távú megőrzés egyetlen útja, ha a történelemben a szöveghordozók tartóssága úgy csökken, ahogy példányszámuk nő. A megőrzésnek további feltétele, hogy az emberek használják is a rendelkezésükre bocsátott nemzeti kultúrát. Olyan – a gyűjtemény bejárását megkönnyítő – rendezéseket kell a felhasználók kezébe adni, amelyek nem a kiadásokat vagy példányokat írják le, mint egy könyvtári katalógus, hanem a szövegeket. A „digitális filológia” a hagyományos filológia rendezési elveit követi, amikor a műveket nyilvántartja, kiadja és narratívába szervezi. A régi magyar versek online nyilvántartása: 1994, az első online kritikai forráskiadás: 1994, az első online, többforrású kritikai kiadás: 1998, az első online irodalomtörténet: 2006. A napjainkban kialakult „digitális bölcsészet” részint a már korábban, gyakran merőben más célból létrehozott digitális eszközök segítségével állapít meg új összefüggéseket. Ezek olykor lényegesek, olykor nem. Az 1991-ben megjelent *A vers* már előrevetítette az anyagismeretről lemondó, nem olvasó filológus lehetőségét, melyre Franco Moretti is felhívja a figyelmet.

Kulcsszavak:

virtuális könyvtár, nemzeti könyvtár, digitális filológia

A 2016-os debreceni Networkshop konferencia vitaforum (2016. március 30-án) a „Digitális bölcsészet és a könyvtárak” címet kapta. Kokas Károly vitaindító előadása után a korreferátumokat szabadon folyó vita követte. Horváth Iván a hozzászólását írásban küldte el. A szerző a dolgozatát eredetileg Monok István születésnapjára (MONOKgraphia, 2016) szánta, de teljes terjedelmében csak itt és most jelenik meg.

A tanulmány megírásának egyik fontos kontextusa az volt, hogy készültek a szerző szakmailag is erősen kötődött (digitalizálási tanácsadóként) a Nemzeti Könyvtárhoz, különösen annak akkor felmerült költözéséhez és a napirendre kerülő nagy volumenű állomány digitalizálásához. Az a megfontolás, hogy a régi értelemben vett központi, nemzeti könyvtár modellje a digitális világban aligha folytatódik majd, föltétlen érvényes megfigyelés. Így az írás jövőképe ma talán még távolabbra mutat, mint készültek, 2016-ban.

Az, hogy az újabb korok könyvtára egészen más lehetőséget biztosít majd a filológia és művelője számára, mint a régiek, s hogy ez messzebbre mutat, hogysem egyszerű módszertani változásnak tekintjük, valójában az előzőekből következő megállapítása a tanulmánynak.

Mivel Horváth Iván és projektjei több, mint negyven éve részesei egy olyan folyamatnak, amelyek a digitális bölcsészet megalapításához, ha tetszik folyóiratunkhoz vezettek, örömmel adunk helyet a szerző visszaemlékező, gondolatgazdag írásának.

A szerkesztők



Egy régi metafora

„Az univerzumot (amelyet mások Könyvtárnak neveznek) meghatározatlan és talán végtelen számú, hatszög alakú galéria alkotja, melyeknek közepén alacsony korláttal körülvevő nagy szellőzőaknáknak vannak. Minden hatszögből láthatók lefelé és felfelé az emeletek – sehol sincs végük.”¹ Hasonlít-e vajon Borges képzeletbeli könyvtára arra a virtuális nemzeti könyvtárra, amelyet az Országos Széchényi Könyvtárban (OSZK) szeretnének létrehozni?

A virtuális könyvtár könyvei

Az egyszerűség kedvéért ne törődjünk most az OSZK kép-, térkép-, hang- és filmállományaival. Maradjunk a szöveges dokumentumoknál, lényegében a – nevezzük így őket – könyveknél. A könyvtári dokumentum jelentéstartománya a „könyv” jelentése körül szóródik, ha a könyvvel szemben nem állítunk fel sem terjedelmi, sem tartóssági, sem a sorozatosságot korlátozó előírásokat. A könyv pedig: a szöveges információ megőrzésének és terjesztésének szokásos eszköze, amely régen agyagtábla, papirusztekercs, pergamenkódex, nyomtatott kiadvány volt, ma pedig főleg weblap.

Olyasféle a helyzet, mint 1500 körül volt. Akkor még bőven készültek díszes kódexek, s a könyvtárak még főleg ezeket gyűjtötték. Voltak ősnymtatványok, amelyek nevetségesen utánózták a kódexeket (rövidítésjelek, ligatúrák, iniciálék, a címlap hiánya, méret és kötés). És voltak már mai jellegű könyvek: Aldus Manutius velencei kiadványai. Ma mindezek megfelelőit megtaláljuk. Itt vannak körülöttünk a közelmúlt papírkönyvei – ezt a formát szoktuk meg, a könyvtárak főleg ilyeneket gyűjtenek –, itt az őket utánzó ősnymtatványok, vagyis a PDF-kiadványok és e-bookok, amelyek mindenfélét megőriznek a papírműltből, a nyomdai tükörtől egészen a gépi másolás megnehezítéséig. Gyűjtésüket repozitóriumok végzik. Végül itt vannak a korszerű kiadványok, a weblapok. Gyűjtésükhöz a legtöbb könyvtárban ma még hozzá sem fogtak. (Majd az önműködő frissítést megkönnyítő könyvtári szabványokhoz való önkéntes csatlakozás lesz az ára annak, ha valaki a kulturális örökség részévé óhajtja megtenni az általa létrehozott weblapot.)

A weblapokat gyűjtő, rendszerező és szolgáltató könyvtár hamarosan általános jelenség lesz. Minden közgyűjtemény szükségképpen könyvtárrá válik, a múzeum és levéltár is, amint digitalizálja (vagyis szöveggé alakítja át), és a hálózaton szolgáltatja anyagát. Még a könyvkiadó is könyvtár attól a pillanattól fogva, hogy megszűnik árusítani könyveit. 16 éve, amikor a *Bábel*-könyvben ismertettem azt a látomást,² hogy a könyvkiadás helyébe könyvszolgáltatás lép, a jövőről beszéltem, de ma ez a jelen. Az Európai Bíróság úgy döntött, hogy a hálózati könyvkiadó nem árusít, hanem szolgáltat. A köteles példány kérdését új módon kell megközelíteni.

¹ Jorge Luis Borges, *Körkörös romok*, ford. Boglár Lajos, Kozmosz fantasztikus könyvek (Budapest: Móra Ferenc Könyvkiadó, 1972), 32.

² Horváth Iván, *Magyarok Bábelben* (JATEPress-Gépeskönyv, 2000), <http://magyar-irodalom.elte.hu/babel/>.

Végtelen-e a bábeli könyvtár?

A bábeli könyvtár annyiban végtelennek tekinthető – és bármely nemzeti könyvtár ugyanilyen –, hogy sohasem telik meg. Minden nemzeti könyvtár bővíthető, minden katalógusba be lehet csúsztatni további tételt. A nagy könyvtárak fejlődése elvileg nem kell, hogy véget érjen. A valóságban azonban véges a nemzeti könyvtár is, a bábeli könyvtár is. Az utóbbiban hiába elég egy könyv létezéséhez annak pusztán lehetősége, hiszen épp a könyvek lehetősége az, ami a bábeli könyvtárban nem végtelen. Meg van adva terjedelmük betűkben, sorokban, lapszámokban, és meg van adva a jelkészlet – nagyon korlátozott – elemszáma. Ebből könnyen kiszámítható az összes kombináció, a bábeli könyvtár összes lehetséges könyvének száma, persze ha a több példányban előforduló, azonos szövegű könyvektől eltekintünk. Óriási szám, de nem végtelen, hanem 1956×10^{1834097} .³ Borges elbeszélője – aki könyvtárát végtelennek hiszi – ezen a ponton avatkozik be hiedelmével. A bábeli könyvtár az elbeszélő szerint végtelen – de csak azért, mert szerinte minden egyes könyvből önmagában is végtelen példány áll rendelkezésre. A megoldás talán nem túl elegáns, de az elbeszélő ehhez folyamodik:

Akik pedig határtalannak képzelik, elfelejtik, hogy a könyvek lehetséges száma nem végtelen. E régi problémára én ezt a megoldást bátorkodom sugallni: *a Könyvtár határtalan és periódikus*. Ha egy örökké élő utazó bármely irányban átutazna rajta, évszázadok múltán meggyőződhetne róla, hogy egyazon kötetek ismétlődnek egyazon rendetlenségben (amely így, ismétlődve renddél, a Renddél válik).⁴

Az író és matematikus Guillermo Martínez két ponton is átalakítja Borges elgondolását, a cantori végtelen halmazokhoz közelítve azt. Először felszámolja a jelkészlet és a könyvterjedelem korlátozását – ezzel most ne törődjünk. Aztán arra hivatkozik, hogy Borgest lenyűgözte a halmazelméletben az a képtelen lehetőség, amikor az egész nem nagyobb, mint egy része.⁵ Ez valóban lenyűgöző. Georg Cantor úgy határozta meg a végtelen halmazt, hogy a halmaz és egy valódi részhalmaza között kölcsönösen egyértelmű megfeleltetés létesíthető.⁶ Vagyis a végtelen halmaz ugyanakkora, mint egy valódi részhalmaza. Én David Hilbert szálloda-példáján értettem meg a dolgot. A végtelen szobát tartalmazó szállodában végtelen utas lakik, minden szoba foglalt. Megérkezik egy turistabusz végtelen sok új utassal. A szállodaigazgató nem esik kétségbe, hanem más szobákba költözteti át korábbi vendégeit. Az 1. szoba lakóját a 2.-ba, a 2.-ét a 4.-be, a 3.-ét a 6. szobába, és így tovább, mindenkit a soron következő páros számú szobába helyez át. Ezzel megürül az 1., a 3., az 5. stb. szoba, vagyis az összes páratlan számú. Ezekben végre elhelyezheti új vendégeit. És ha végtelen sok busz jön, mindegyikben végtelen utassal, ők is mindannyian elhelyezhetők a páratlan számú

³ Wikipedia, „La biblioteca de Babel,” hozzáférés: 2018.05.15, https://es.wikipedia.org/wiki/La_biblioteca_de_Babel.

⁴ Borges, „Körkörös,” 42.

⁵ Guillermo Martínez, *Borges és a matematika*, ford. Kutasy Mercédesz (Budapest: Európa Könyvkiadó, 2010), 19.

⁶ Ian Stewart, *A matematika problémái*, ford. Freud Róbert (Budapest: Akadémiai Kiadó, 1991), 74.

szobákban, mert páratlan számból is ugyanolyan sok van, mint ahány természetes szám.⁷

Martínez a végtelen halmaz (a végtelen sok szobából álló szálloda) és annak ugyan- csak végtelen részhalmaza (a páratlan számú szobák) közötti megfeleltetést a novella végén olvasható lábjegyzetben találja meg:

Letizia Álvarez de Toledo megfigyelte, hogy a hatalmas Könyvtár felesleges; tulajdonképpen elegendő volna *egyetlen kötet* szokványos formátumban, kilenc- vagy tízpontos betűből szedve, amely kötet végtelen számú, végtelenül finom lapokból állna. (Cavalieri a XVII. század kezdetén azt mondta, hogy minden szilárd test végtelen számú egymásra helyezett sík.) E selymesen finom 'zsebkalauz' kezelése nem lenne kényelmes; minden látható lap ugyanolyan lapokká válna szét; a megfoghatatlan középső lapnak nem lenne hátsó oldala.⁸

Borges a *Homokkönyv* című novellájában ezt a könyvet írta meg, teszi hozzá Martínez.

Elismerem, hogy Martíneznek igaza lehet abban, hogy Borgest a halmazelmélet, közelebbről a végtelen halmazoknak ez a fantasztikus tulajdonsága ihlette meg, de a párhuzam szerintem nem kielégítő. A cantori meghatározás értelmében a megfeleltetésnek a végtelen halmaz és annak egy valódi részhalmaza között kellene megtörténnie, de Letizia Álvarez de Toledo képzeletbeli könyve nem része a bábeli könyvtárnak, nem annak egy polcán foglal helyet. (A Homokkönyv végül könyvtárba kerül, de szokványos méretűbe.)

Martíneznél az *Alef* című novella is kapcsolatba kerül a végtelen halmazokkal. A végtelen halmaz a világ, valódi részhalmaza pedig az Alef, amely az egész világmin- denségről, a világ legkisebb eleméről is tartalmaz információt. Így fennáll a halmaz és valódi részhalmaza közötti szigorú megfeleltetés – hiszen az Alef is része a világnak –, de ez nem könyvtári história.

Borges 1941-ben közölte a *bábeli könyvtárt*, weblapokról még nem tudhatott. Martínez azonban talán felvethette volna a lehetőséget. A hiperhivatkozásra és a beágya- zásra való képesség a weblapoknak nagyon érdekes tulajdonságai.

Vegyünk például egy weblapokból álló képzeletbeli könyvtárat. A weblapok között mindenféle lehet, de mindegyik formailag hasonló metaadatokkal felszerelt, hasonló kinézetű címlappal kezdődik – elvégre mindegyikük könyvtári egység –, és mind- egyikük a könyvtár valamelyik tartalmára való rámutatással végződik. Ez a tartalom lehet digitalizált könyv, digitális papírkönyv-utánpótlás, és lehet, persze, eleve weblap. Az egyik weblap azonban különleges kötet: ez a könyvtár katalógusa. Olyasféle címlappal kezdődik, mint a többi (mi is ez, ki csinálta, mikor stb.). A tartalma azonban a könyvtár többi kötetének metaadatait sorolja fel, hiszen katalógus. Ma már elvárható lenne egy weblap-alapú katalógustól, hogy a benne leírt kötetek tartalmát is – beágyazás vagy hiperhivatkozás révén – egyetlen egérkattintással előhozza. Fennáll a teljes halmaz (a könyvtár) és részhalmaza (a katalógus) közötti megfelelés. A megfelelés annyira szo- ros, hogy igazából meg sem tudjuk különböztetni a katalógust a könyvtártól magától.

⁷ Stewart, *A matematika problémái*, 70.

⁸ Borges, „Körkörös,” 43.

Ha a könyvtárban megvan a saját katalógusa, az azt jelenti, hogy a könyvtár az egyéb könyvtári tételek között önmagát is tartalmazza.

A nagy, hagyományos könyvtárak gyűjtőkörébe a többi nagy könyvtárak nyomtatott katalógusai is beletartoznak. Nincs akadálya, hogy képzeletbeli weblap-könyvtárunkba akár tucatjával besoroljuk más könyvtárak webkatalógusait, természetesen az illető könyvtárak anyagaira mutató hiperhivatkozások özönével egyetemben. Könyvtárunk ettől kezdve jóval nagyobb, mint ő maga.

Persze nem végtelen. A valóságos tárgyak között talán nincs is végtelen. A világegyetem összes elemi részecskéinek száma bizonyára nagyon nagy szám, de nem végtelen, sőt talán ugyanolyan messze esik a végtelentől, mint az egy. De egy nagy weblap-könyvtár, amely hasonlít a végtelen halmazhoz abban a tulajdonságában, hogy vannak olyan valódi részhalmazai, amelyek nem kisebbek, mint ő maga, felveti azt a kérdést, hogy hová tegyük, fizikailag hol helyezzük el ezeket a részhalmazokat.

Építészet

Gyönyörű könyvtár volt az OSZK, amelyben ifjúkoromat töltöttem. A Magyar Nemzeti Múzeum északi szárnyában helyezkedett el. Bejárata a Bródy Sándor utcáról nyílt. Át kellett vágni a római kőtár félhomályos folyosóján, aztán föl az első emeletre, a dohányzóban az ember elszívott egy búcsúcígit, majd a hosszú katalógusfolyosó végén nem egyenesen a hírlapolvasóba, vagy jobbra, a nagy olvasóterembe lépett be, hanem balra fordult, és rövid katalógusfolyosó után végre megérkezett a Régi és Ritka Nyomtatványok Tárába és a Kézirattárba. A Könyvtár 1957-ben eltávolított vezetője, Varjas Béla hibának tartotta a Várba költözést. Olyan épülethez ragaszkodott volna, amely eleve könyvtári célra épült. Az irdatlan palotaszárny végül nem lett használhatatlan, inkább vicces. A pöffeszkedő épület dölőfét enyhítik az oda nem illő pultok és katalógusok. A két óriási Atlaszon határozottan javít, hogy megfeszített kőizmaikkal nem tartanak semmit. Most viszont helyreállítanák a Ferenc József- és Horthy-állapotot, megismételve és így négyzetre emelve a giccset. Könyvtári munka az új esztétikai környezetben reménytelen lesz. El kell költözni.

Azt hittem, az én ötletem volt, hogy a Magyar Rádió telephelyére, a Magyar Nemzeti Múzeum szomszédságába kerüljön vissza az OSZK. Rácz Ágnes azonban kiderítette, hogy az OSZK eleve arrafelé terjeszkedett volna, és a szóban forgó telkeket már az 1930-as években megszerezte.

Az építészeti tervezéshez szeretnék két elemi szempontot javasolni. (Addig kell szólni, amíg lehet. A Francia Nemzeti Könyvtár építésze üvegtornyokba rakta a fénytől és hőtől védendő könyveket, viszont alagsorba száműzte az olvasókat.)

Több, mint 3500 OSZK

Az OSZK egyfelől azzal emelkedik ki a többi magyar könyvtár közül, hogy messze a legnagyobb, másfelől azzal, hogy elvileg teljességre törekvő a hungaricum-gyűjteménye. A digitalizálás után tételei weblapokká válnak, weblap-könyvtárak között pedig bármely méretkülönbség könnyen az ellenkezőjére fordul. Akármelyik weblap-könyvtár – a közművelődési könyvtár is – beágyazhatja vagy hiperhivatkozhatja a teljes hungaricum-gyűjtemény weblapjait. Új képességének hála a köz-

művelődési könyvtár forradalmi átalakuláson megy át. A változást építészeti – főleg belsőépítészeti – eszközökkel látványosan kifejezésre lehet juttatni.

Mivel minden közművelődési könyvtár teljes értékű virtuális OSZK-vá változtatható, helytelen lenne, ha olvasótermük felszerelése különböznék az OSZK általános olvasótermeitől. Ha jól látom, valamivel több, mint 3500 települési könyvtár található Magyarországon.⁹ Az építész feladata az, hogy az OSZK olvasótermeinek csak csekély részét építse fel a főváros VIII. kerületében. 3500 OSZK-olvasótermet az ország különböző településein kell létrehozni.

Talán a közművelődési könyvtárosok is az OSZK alkalmazásába kerülnek. A pákozdi művelődési házban – nyaraláskor az ő könyvtárukat szoktam használni – az OSZK egyik teljes értékű olvasótermébe léphetnék be. Az olvasóterem pedig olvasókörré válnék.

Az új Magyarország egyetlen évszázad alatt létrejött. A felvilágosodás eszméit 1770 körül kezdték hazahordani a testőrírók, Budapest pedig 1870 körül már kezdett világvárossá válni. Magyarország megújítói mai mércével elképesztően sokat és sok nyelven olvastak. A többségében külföldi tudományos és az eleinte főleg idegen nyelvű szépirodalom beszerzése és értelmezése meghaladta az egyén képességeit, ezért a művelődni vágyók olvasóegyletekbe (*Lesekabinett*) tömörültek, s ott beszéltek meg az elolvasottakat. Eredetileg olvasóegylet-jellege volt a Széchenyi alapította Pesti (később Nemzeti) Casinónak, sőt a jobb kávézóknak is. Kávéház-látogatóként Petőfi Pestről figyelemmel tudta kísérni az egykorú párizsi sajtót.

Olvasmányokhoz hozzájutni ma sokkal egyszerűbb, mint Petőfi korában, de az olvasmányok megértése – pláne idegen nyelveken – semmivel sem könnyebb. A nemzeti és nemzetközi szövegörökség és a világháló friss olvasmányözönében újra olvasó- és internetező körök jönnek létre, hogy az olvasók segíthessék egymást. Olvasókörré fog átalakulni az OSZK több ezer olvasóterme, a székhelytől a közművelődési telephelyekig. A könyvtárban régebben csend honolt. Az olvasókör a halk beszélgetése. A meglevő nagy olvasótermeket vékony, de hangszigetelő falakkal kisebb-nagyobb fülkékre fogják szabdalni. A felhasználói felületek elé olyan olvasók telepednek, akik félhangosan megosztják egymással észrevételeiket. A felhasználók nagy, kerek üvegasztalokat ülnek körül, körülbelül akkorákat, mint a Petőfi Irodalmi Múzeumban őrzött márvány Pilvax-asztal. Az üvegasztalokon számos sugárirányú, szabadon méretezhető képernyőablak elfér, de kávésbögrét is rájuk szabad tenni. Az olvasók közti beszélgetés engedélyezése harmonizál a világháló interaktivitásával. A könyvtáros a beszélgetést, az olvasótársak közötti együttműködést szakértelmével segíti. Az együttműködés nemcsak művelődési, hanem bármiféle célt szolgálhat, akár üzletit is. A hálózati olvasókörré átalakuló közművelődési könyvtár virtuális tartalomipari park. Vagy iskola, amelyben az ismeretszerzés tempóját a tanulók maguk szabályozzák.

⁹ „Települési könyvtárak (2000–),” Központi Statisztikai Hivatal, hozzáférés: 2016.03.16, https://web.archive.org/web/20160316064317/https://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_zkk003.html.

Egyetlenegy OSZK

Régebben annak, aki szöveget akart olvasni, szöveghordozó dokumentumot kellett a kezébe vennie. A könyvtári katalógust böngészve eljutott a dokumentumig, majd abban a szövegig. A digitalizálás után a szövegek immár a hordozó dokumentumokról leválasztva is hozzáférhetők. A szövegigény közvetlenebb módon elégül ki.

De vajon hiteles-e a szöveg? Erre a kérdésre az átirás ellenőrzése a válasz. A digitális szövegek korában az ellenőrző könyvtárlátogató nem a régi módon, a dokumentumtól jut el a szöveghez, hanem az otthon is rendelkezésére álló szövegtől jut el a dokumentumhoz. Az ellenőrzés vagy egyáltalán bármilyen elmélyült szöveg tanulmányozás során az eredeti dokumentum – egy ritka kiadás, egy szerzői kézirat – kézbe vétele is elkerülhetetlen. Az OSZK ebből a szempontból megőrzi kivételes jellegét, mint a szöveges hungaricumokat hordozó könyvtári dokumentumok elvi teljességre törekvő gyűjteménye. Tüske László ezért szereti az OSZK-t új *hiteles helynek* nevezni. A bárhol, akárhány példányban lemásolható, virtuális OSZK így értékeli föl a régi szöveghordozók egyedülálló lelőhelyét.

Ezt a körülményt is kifejezheti az építészet. Az OSZK-nak átadandó, rádiós telkeken van egy igen érdekes épület, a hidegháború figyelemre méltó emléke. Az 1952-ben, roppant vastag vasbeton falakból emelt fedezéknek két földalatti és két földfelszíni szintje van. A bennük egykor elhelyezett stúdióknak elvben atomcsapás után is meg kellett volna őrizniük működőképességüket. A betonkocka köré Nánási Sándor és Szende László tervezett megnyugtatóan almazöld irodaépületet (1969), az atombunker mibenlétét díszkő-borítással és Somogyi József szoborcsoportjával – sikertelenül – leplezve.¹⁰

Az építészeti tervnek valamiképpen ki kellene fejeznie az OSZK egyedülálló dokumentum-megőrző-, kincstár- és hiteleshely-jellegét, azt, hogy ez a szövegörökségre vonatkozó viták eldöntésének végső instanciája. Ahol a szövegörökség szövegeit eredeti hordozójukon, bármely kiadásban, nemritkán kéziratban kézbe lehet venni. Talán ez a történelmi levegőt árasztó, furcsa doboz is valami módon beilleszthető lenne a felfogásba. Hiszen a nyelvemlék-kódexeket, a Corvinákat, a magyar antikvákat és klasszikusaink kézírathagyatékát csakugyan atom- és üstökösálló kincstárban kellene tartani.

A következőkben azt vizsgálom meg, hogy a digitális bölcsészetet mennyire lehet használni a virtuális nemzeti könyvtár létrehozásában, működtetésében és megőrzésében.

A régi vágású digitális filológia a könyvtárban

A *Digital Philology* folyóirat főleg azoké,¹¹ akik a középkor-kutatásban szívesen vesznek igénybe új eszközöket, és akik közül az öregebbek valaha „új filológusoknak” ne-

¹⁰ Leguan, „Magyar Rádió épületei, Esterházy és Károlyi palota,” 2014. ápr. 25., <http://egykor.hu/budapest-viii--kerulet/magyar-radio-epuletei/3895>.

¹¹ Editorial Board of *Digital Philology*, hozzáférés: 2018.05.15, https://www.press.jhu.edu/journals/digital_philology/editorial.html.

vezték magukat. Nemzetközi konferenciát is tartottak Budapesten, 2000-ben.¹² Munkájuknak nincs különösebb könyvtári vonatkozása. Irányítsuk figyelmünket őhelyetük most azokra a számítógépes filológusokra, akiknek a digitális örökség megőrzésében – legalábbis szerintem – kulcsszerepük lesz. Elsősorban az ő felelősségük, hogy a könyvtárban megőrzött múlt ne maradjon megműveletlenül.

A digitális állományok hosszú távú megőrzése gyakori könyvtári téma.¹³ Mit kell tenni, ha megváltozik a gépi környezet, a hordozó, a szoftver, a betűkódtábla? Mit tegyünk, ha nyugdíjba megy a bácsi, aki még tudott programozni a Micro CDS/ISIS adatbázis Pascal nyelvjárásában? Ha ügyesek vagyunk, akár évtizedekig életben tarthatunk digitális dokumentumokat. A hosszú távú megőrzés valójában igencsak rövid időtartamot jelent. A ninivei könyvtárosok furcsállnák is, hogy ma ilyesmit nevezünk hosszú távú megőrzésnek. Asszurbanipál könyvtárából másfél ezer agyagtábla még ma is megvan – pedig azokat az i. e. 7. században gyűjtötték, rendszerezték, szolgáltatták. Miként tudja az irodalomtudomány segíteni a könyvtárost abban, hogy nemzedékeken keresztül fennmaradjanak a digitális nemzeti könyvtár dokumentumai?

Információtörténeti szabály, hogy a szöveghordozók tartóssága az ókor óta egyre csökken, példányszámuk viszont nő. A digitális könyvtári tartalom hordozóira a szöveget olyan csekély energiával rögzítik, hogy az információ letörléséhez is alig kell energiát igénybe venni. Az ilyen tartalom a jelenlegi csekély példányszámban (a bolygó néhány tucatnyi szerverparkján) nem képes tartósan – évezredekig – fennmaradni. Ha a jövőre nézve is érvényesnek fogadjuk el a történelemből levont szabályt, akkor a digitális könyvtári tartalom megőrzésének egyetlen lehetősége az, hogy szétosztjuk az egészet. A virtuális nemzeti könyvtárat több millió példányban lemásoljuk, és az embereket rábírjuk arra, hogy taneszközként az otthonukban tartsák, vagy mulatság kedvéért vigyék magukkal a kézitáskájukban vagy a zsebükben.

2004-es OSZK-előadásomban még nem tudtam megmondani, hogy milyen műszaki megoldásoktól remélhetjük ezeket a kis méretű és energiaigényű másolati példányokat. Keszei Ernő 2014 őszén az OSZK-ban tartott termodinamikai-információelméleti előadás-sorozata szerint a nagy energiájú lézerrel mélységben írt kvarckristálytól és a betűhordozó DNS-től. De miként tudjuk rátukmálni az olvasókra az irdatlan kulturális örökséget? Csakis narratívákkal. És már az ún. *kulturális örökség gép*¹⁴ létrehozása előtt, már ma is.

A régi könyvtárban a katalóguson keresztül lehetett hozzájutni a dokumentumhoz, azzal pedig a dokumentum által hordozott szöveghez. A mai felhasználónak nincs szüksége a szöveghordozó dokumentumpéldányra, így könyvtári katalógusra sincs. A szöveghez vad keresésekkel, kulcsszavakkal, nevekkkel is eljuthat, de igénybe vehet előrerendezéseket is. Ehhez narratívákban, mesékben feldolgozott digitális örökség kell. Ha nem csinálunk okos rendezéseket a különböző műveltségi szintű felhasználók számára az elkülönített cigányiskola önművelésre kárhoztatott tanulójától a pamlagán

¹² Seláf Levente, szerk., *Vers une nouvelle philologie* (Budapest: Hallgatói Információs Központ, 2007), <http://www.tankonyvtar.hu/hu/tartalom/tkt/vers-une-nouvelle/adatok.html>.

¹³ *Wikipedia*, „Digital preservation,” hozzáférés: 2018.05.15, https://en.wikipedia.org/wiki/Digital_preservation.

¹⁴ Horváth Iván, „Kulturális Örökség Gép,” in Horváth Iván, *Gépeskönyv* (Budapest: Balassi, 2006), 237–248.

regényt olvasó mérnökig és a vizsgára készülő kényszerolvasóig, a virtuális könyvtár anyagát kevesen fogják igénybe venni.

A katalógust elkerülő szöveghezáférést segíti a legegyszerűbb rendezés, a bibliográfia. Az OSZK-ban és Szegeden fogyasztói szemléletű, a felsőoktatásban is remekül használható bibliográfiai szakadatbázis áll az olvasók rendelkezésére, a HUMANUS.¹⁵ Könyvtárban készül, de gyűjtőköre könyvtárközi. Könyvtárközi – és sokkal több bibliográfiánál – az OSZK nyelvemlék-honlapja is,¹⁶ amely példamutató módon kétféle – tudományos és ismeretterjesztő – vetületben tárja fel és aktivizálja a felhasználó számára a legértékesebb levéltári- és kódexállományt.

Könyvtárfüggetlen, egyértelműen szövegközpontú rendezési elvet követ minden valódi hálózati kritikai kiadás. (Valódi akkor, ha nem pusztán egy korábbi papírkiadvány internetes átirata, hanem a legjobb kézirat és nyomtatott források egybevetésével önállóan megállapított és jegyzetelt kritikai szöveg, és eleve weblapnak készült.) A weblap a legnagyobb természetességgel vezeti el, mondjuk, a *Nemzeti dal* olvasóját a virtuális könyvtár mélyére, a költemény legmegbízhatóbb kézirati és nyomtatott forrásainak betűhű átirataiig, sőt fénykép-hasonmásáig, de zenei vagy mozgóképes lehetőségei is korlátlanok. A hálózati kritikai kiadás benyomásom szerint szinte magyar műfaj. Csak azt említem, ami véletlenül nem került el a figyelmemet. Ide tartozik M. Nagy Ilonának és munkatársainak *Margit-legendája*,¹⁷ a néhai Bárcki Ildikó műhelyében elkezdett¹⁸ és ma is sok ágon folytatódó¹⁹ *Érdy-kódex*-kiadás, az egykori Balassi-kiadás,²⁰ Parádi Andrea *Árgirus*-kiadása,²¹ Tóth Tünde Rimay-kiadása²² és Földes Zsuzsával közös *Fanchali Jób-kódex*,²³ Debreczeni Attilának és tanítványainak szövegbázisa,²⁴ különösen Kazinczy művei,²⁵ Devescovi Balázs Eötvös-

¹⁵ Humántudományi tanulmányok és cikkek adatbázisa (HUMANUS), hozzáférés: 2018.05.15, <http://www.oszk.hu/humanus/>, <http://humanus.bibl.u-szeged.hu/>.

¹⁶ Magyar Nyelvemlékek, Országos Széchényi Könyvtár, 2009–2018, hozzáférés: 2018.05.15, <http://nyelvemlekek.oszk.hu/>.

¹⁷ M. Nagy Ilona, szerk., *Margit-legendája: Szent Margit élete, 1510 és latin forrásai* (deba.unideb.hu, 2014), http://deba.unideb.hu/deba/Margit-legendája_Szent_Margit_elete_1510/index.html.

¹⁸ Az Érdy-kódex Munkacsoport honlapja, hozzáférés: 2018.05.15, <http://magyar-irodalom.elte.hu/erdy-kodex/>.

¹⁹ A Sermones Compilati Kutatócsoport Érdy-kódex kiadásának honlapja, hozzáférés: 2018.05.15, <http://sermones.elte.hu/erdy>.

²⁰ Horváth Iván és Tóth Tünde, szerk., *Balassi Bálint összes verse, hálózati kritikai kiadás* (Budapest: Gépeskönyv, 1993–2002), <http://magyar-irodalom.elte.hu/gepesk/bbom/cimlap.htm>.

²¹ Parádi Andrea, szerk., *Árgirus históriája, internetes kritikai kiadás* (Budapest: Gépeskönyv, 2005), <http://www.tankonyvtar.hu/hu/tartalom/tkt/argirus-historiaja/adatok.html>.

²² Tóth Tünde, szerk., *Rimay János ifjúkori versgyűjteménye* (Budapest: ELTE BTK BIÖP–Bibliopolisz, 2007), <http://www.tankonyvtar.hu/hu/tartalom/tkt/rimay-janos-ifjukori/adatok.html>.

²³ Földes Zsuzsa és Tóth Tünde, szerk., *Fanchali Jób-kódex* (Mezura–ELTE BTK CHER–Bibliopolisz, 2008), <http://www.tankonyvtar.hu/hu/tartalom/tkt/fanchali-job-kodex/adatok.html>.

²⁴ Az MTA–DE Klasszikus Magyar Irodalmi Textológiai Kutatócsoport honlapja, hozzáférés: 2018.05.15, <http://textologia.unideb.hu/>.

²⁵ Debreczeni Attila, szerk., *Kazinczy Ferenc összes művei* (Debreceni Egyetemi Kiadó, 2016–2018), http://doi.org/10.0.21.108/Kazinczy_Ferenc_muvei.

regénye,²⁶ Kassák lapjának, *A Tett* folyóirat közzététele,²⁷ Bengi László és Parádi Andrea *Aranysárkány*-kiadása,²⁸ az elsietett,²⁹ majd javított³⁰ József Attila-próza-kötet. A Palkó Gábor által vezetett PIM-projekt, a Digiphil honlapján a papírkidatással párhuzamosan vagy annak nyomán készült, de teljes értékű Mikszáth-weblapok Hajdu Péter³¹ és Labádi Gergely³² munkái; félig-meddig az *Aranysárkány* is ide sorolható. Módszertanilag roppant változatos, gazdag anyag, amely a közeljövőben – a debreceni műhely³³ és a Digiphil becsvagyó tervei³⁴ következtében – gyorsan tovább fog növekedni.

Az irodalomtudományi adatbázisoknak se szeri, se száma. Talán a legújabb Bartók Zsófia Ágnes exemplumtára,³⁵ a legöregebb pedig bizonyára az 1601 előtti magyar versek sok szempontú, teljes nyilvántartása, az RPHA,³⁶ amelyhez most sokkal igényesebb 17. századi kiegészítés készül.³⁷ Elmesélek egy érdekes kísérletet arról, hogy miként lehet a szövegyilvántartásokat beállítani narratívákba.

Az RPHA készítésekor a középkori és 16. századi versek minden előfordulását gondosan elemeztük. A szakszeminárium hallgatói hozzám jöttek, amikor elakadtak, mert kritikai kiadás híján nem tudták, hogy a szóban forgó vers melyik változata az „eredeti”, melyik alapján kell megállapítani a terjedelmet, a szerzőt, a versformát – mert mindez változékonynak bizonyult. Egy ideig rá-rávtam a választ, olykor csak megérzésből, rossz lelkiismerettel, aztán, néhány év után, talán 1979 körül nagy levegőt vettünk, és kimondtuk, hogy nincs eredeti, hanem csak változatok vannak. Az RPHA adatmodellje végül ezt tükrözi. Abból indultunk ki, hogy egy-egy vershez több

²⁶ Devescovi Balázs, szerk., *B. Eötvös József: A' falu' jegyzője, 1845* (Budapest: ELTE BTK BIÖP, 2003), <http://magyar-irodalom.elte.hu/totnes/eotvos/>.

²⁷ *A Tett* (2014), <http://digiphil.hu/context:atett>.

²⁸ Bengi László és Parádi Andrea, kiad., Parádi Andrea, szerk., *Kosztolányi Dezső: Aranysárkány* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:aranysarkany>.

²⁹ Horváth Iván, főszerk., *József Attila összes tanulmánya és cikke* (Gépeskönyv, 1999–2001), <http://magyar-irodalom.elte.hu/ja/cimlap.htm>.

³⁰ Horváth Iván és Fuchs Anna, szerk., *József Attila összes tanulmánya és cikke: Szövegek, 1930–1937* (Gépeskönyv, 2012), <http://textus.elte.hu/ja/>.

³¹ *Mikszáth Kálmán összes művei, 42* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:mkk42>.

³² *Mikszáth Kálmán összes művei, 39* (Budapest: Petőfi Irodalmi Múzeum, 2014), <http://digiphil.hu/context:mkk39>.

³³ Bodrogi Ferenc Máté, „Elektronikus kritikai kiadások a klasszikus magyar irodalomban,” előadás a *Számítógép az irodalomtudományban* c. konferencián, 2015. november 24., http://videotorium.hu/hu/recordings/details/12382,Elektronikus_kritikai_kiadasok_a_klasszikus_magyar_irodalomban.

³⁴ Palkó Gábor, „Mit jelent a digitális filológia a szemantikus web korában?” előadás a *Számítógép az irodalomtudományban* c. konferencián, 2015. november 24., https://videotorium.hu/hu/recordings/details/12388,Mit_jelent_a_digitalis_filologia_a_szemantikus_web_koraban.

³⁵ Bartók Zsófia Ágnes, szerk., *Régi Magyar Exemplumadatbázis* (ELTE BTK Régi Magyar Irodalom Tanszék, 2014), hozzáférés: 2018.05.15, <http://sermones.elte.hu/exemplumadatbazis/>.

³⁶ Horváth Iván, főszerk., *Répertoire de la poésie hongroise ancienne* (ELTE–OSZK, 1992–2018), hozzáférés: 2018.05.15, <http://rpha.elte.hu/>.

³⁷ A XVII. századi magyar vers repertórium, OTKA-projekt, 2014–2016, vezető: Szilasi László, hozzáférés: 2018.05.15, <http://nyilvanos.otka-palyazat.hu/index.php?menuid=930&num=109227&keyword=109227>.

nyomtatott vagy kéziratos könyv tartozhat, amelyekben előfordul, egy-egy könyvhöz pedig több vers is, amelyek előfordulnak benne. A *record*, az adatfelvételi egység nem a vers, nem is a könyv, hanem egy vers és egy könyv egyszeri találkozása. Ez mindig pontosan leírható. A mi adatmodellünk hálószerkezetű, nem pedig fa, mint a – platonizmussal nem ok nélkül jellemezhető³⁸ – előkelő könyvtári adatmodell, az FRBR.³⁹ Viszont a mi összes *könyvünk* mindig tökéletesen ellentmondásmentesen megfelel valamely tényleges könyvtári kéziratnak vagy kiadásnak. Mivel az RPHA minden tétele eleve könyvtári, a versnyilvántartás lap- és levélszámai pontosan kijelölik a versek lelőhelyét a digitalizált könyvtári állományokban. A régi könyvek digitalizálása nagyrészt már végbement, a megfelelő program (Király Péter munkája) rég kész, de az RPHA és a digitalizált állományok közötti szoftveres összekapcsolás munkaszervezési okokból csak észbontó lassúsággal történik meg. Mindenesetre amikor 2002-ben – Szegedy-Maszák Mihály felszólítására – megírtam *A magyar irodalom történeteinek próbafejzetét*, akkor úgy tekintettem, mintha az RPHA minden vers-adatlapja már előhívta volna az illető vers minden kéziratos vagy nyomtatott előfordulásának nagy felbontású fényképét. (Mintha már meg is tette volna azt, amire elvileg képes volt.) A próbafejzetben úgy tekintettem át a régi magyar vers verstani és műfaji rendszerét, hogy betartottam néhány titkos játékszabályt.

- I. Nem léptem át a fejezetekre megszabott maximális terjedelmet, a 40 ezer leütést. (A fejezet éppen ennyi.)
- II. Egy esszészerű bevezető után az áttekintés kizárólag az RPHA-nak feltett kérdésekből (lekérdezésekből) állt. Minden válasz a versek darabszámát mondta meg, és így is került be a tanulmányba (milyen fajta vers mennyire volt gyakori).
- III. A darabszám kinyerése után az RPHA természetesen tovább faggatható lett volna. Meg tudta volna adni a versek adatlapjait, rajtuk az irodalomtörténeti és poétikai adatokon felül a bibliográfiaiakat is, amelyek megmutatják, hogy az illető vers mely régi kéziratok és nyomtatványok mely lapjain szerepel. Akkor még az illető lapok könyvtári fénykép-hasonmásaival az összekapcsolás nem történt meg, sőt a digitális fényképek még nem is léteztek. Ezt az utat tehát egyelőre nem jártam végig, de elvileg végigjárhattam volna. Vagyis minden RPHA-válasz esetében a konkrét vers legkonkrétabb előfordulásait is meg lehetett volna kapni, és tudtam, hogy egykor majd eljön az idő, amikor ténylegesen is megkapjuk őket.
- IV. Az irodalomtörténeti narratívába mennél több verset, pontosabban a teljes versállományt hiánytalanul be akartam kapcsolni. Mivel műfaj- és verstan-történetet írtam, e két szempontból tettem fel kérdéseket az egész anyagnak, és mindkét szempontból elvégeztem a teljes anyag mennél finomabb osztályozását.

³⁸ Golden Dániel, „FRBR-álmom és könyvtári valóság,” előadás az Országos Széchényi Könyvtárban, 2015. szeptember 14.

³⁹ IFLA Study Group on the FRBR, *Functional Requirements for Bibliographic Records* (Munich: K. G. Saur Verlag, 1998), <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>.

- V. Az eredmény az, hogy egy 40 ezer leütésnyi fejezetben az 1600. december 31. előttről ránk maradt, magyar nyelvű versanyag minden egyes versét *legalább két, de inkább jóval több fajta* (műfaji és verstani) osztályozásban szerepeltettem.
- VI. Sikerült elérni azt a kívánatos állapotot, amikor az irodalomtörténész meg tudta adni, hogy pontosan mely művekből áll az a halmaz, amelynek történetét megírja.

Úgy képzeltem, hogy a jövő irodalomtörtészének (aki narratívát visz be a holt, könyvtári anyagba) kutyakötelessége e könyvtári anyag lehető teljességét narratívába szervezni. Bezeczky Gábor egy előzetes közleménye,⁴⁰ valamint a Batori Anna – Labádi Gergely szerzőpáros egy előadás-kivonata⁴¹ nyomán nem zárom ki, hogy csakugyan errefelé, a most *distant reading*nek, vagyis *távolsági olvasás*nak nevezett nagy áttekintések felé tartunk.

Az irodalomtörténet, amelybe a próbafejezet készült, két változatban jelent meg. Szegedy-Maszák főszerkesztésével három vaskos, nyomtatott kötet készült,⁴² hallgatóim és én pedig a 2006/07-es tanévtől egy hálózati változatot, a *Villanyспенót*⁴³ szerkesztettük, pontosabban *szerbantalosdit* játszottunk különböző szemináriumokban. A *Villanyспенót* nagyon hasonlított arra, amilyennek Rákai Orsolya ma a soron következő, az Irodalomtudományi Intézetben készülő irodalomtörténeti kézikönyvet láttatja.⁴⁴ Ez is gumiból volt, itt se volt állandó a fejezetek száma, sőt a szövegük sem. Az én egyik fejezetemből például (*Ómagyar szövegemlékek mint textológiai tárgyak*) a buzgó internetező három különböző változatot talál, köztük ma már egy kis könyvet is.⁴⁵ Interaktív volt: a hallgatók és mások telerakták hiperhivatkozásokkal a fejezeteket, sőt vizsgákhoz *puskákat* (tömörítvényeket) is készítettek hozzájuk. A puskák nyilvánosak. A szerzők is elolvashatták őket, és tiltakozhattak a félreértések ellen. Lehetett a fejezeteket bírálni, javítani, sőt szaporítani is. Elég sok új fejezet készült. A *Villanyспенót* a vétőjogig interaktív volt: a szerkesztő-hallgatók számúzták a fejezetek közül a fékevesztett halandzsákat. Most már pár éve magára hagytuk: egy hálózati kézikönyv nem fejlődhet valamilyen intézményes élet nélkül. Eleve nem szántuk többnek, mint afféle labornak a későbbi irodalomtörténetek számára. Ha az OSZK-ban megalakulna a digitális filológiai osztály, még lehetne játszani vele. Kaphatna egy egyetlen képernyőoldalnyi nyitó irodalomtörténetet, afféle narratívapróbát, aztán

⁴⁰ Bezeczky Gábor, „Az időszaki sajtótermékek számának alakulása a 19. századi Magyarországon,” in *Szolgáltatokat ajánlom a 60 éves Jankovics Józsefnek*, szerk. Császtvay Tünde és Nyerges Judit, (Budapest: Balassi-MTA ITI, 2009), 63–69, <http://www.balassikiado.hu/BB/NET/StudioIum/Bezeczky.pdf>.

⁴¹ Batori Anna és Labádi Gergely, „Egy regényadatbázis felépítése – kérdések és lehetőségek,” előadaskivonat a *Számítógép az irodalomtudományban* c. konferenciára, 2015. november 24., <http://dhu2015.mit.bme.hu/program/eloadasok/batori-anna/>.

⁴² Szegedy-Maszák Mihály, főszerk., *A magyar irodalom története* (Budapest: Gondolat, 2007).

⁴³ Horváth Iván, főszerk., *Villanyспенót: A magyar irodalom története*, hálózati kiadás, hozzáférés: 2018.05.15, <https://irodalom.oszk.hu/villanyспенot/>.

⁴⁴ Rákai Orsolya, „Irodalomtörténeti alapfogalmak vagy alapformák?” előadás a *Számítógép az irodalomtudományban* c. konferencián, 2015. november 24., http://videotorium.hu/hu/recording/s/details/12399,Irodalomtorteneti_alapfogalmak_vagy_alapformak_.

⁴⁵ Horváth Iván, *Ómagyar szövegemlékek mint textológiai tárgyak* (Budapest: OSZK, 2015), <http://mek.oszk.hu/15400/15446/>.

más narratívapróbákat – és összelegózhathatnánk különböző irodalomtörténeti meséket belőle. Laborgyakorlat lehetne a készülő intézeti nagy mű és más próbálkozások számára.

Ne tévesszük szem elől a fő célt. Azért írunk, *ma* azért írunk irodalomtörténetet, hogy narratívákat adjunk a digitális örökséggel nyakon öntött felhasználónak, hogy az adófizető közösség tagjaira – mindenkire – rászakadó irdatlan kultúra ne legyen elrémitő. Ehhez „Szerb Antal”-t kell írni, nem „Nagy Pintér”-t, nem „Spenót”-ot. Az eddigi hivatalos irodalomtörténeteket csak a kényszerolvasók és a hivatásosok búvárolják, Szerb Antalt viszont minden irodalomkedvelő. Nem csak szellemessége miatt. Módszertanilag is sokkal világosabb és egyszerűbb, mint a hivatalos irodalomtörténetek voltak.⁴⁶ A sokszerzős irodalomtörténetek szerkesztői óhatatlanul egyeztetni kényszerülnek, mint a politikusok, pedig a tudomány és a módszer nem többségi döntés kérdése. (Közömbös, hogy hányan vélik úgy, hogy a Föld lapos.) Az egyeztetések végére érve a hivatalos irodalomtörténet szerkesztője elérheti, hogy ne más, hanem ő beszéljen, Szerb Antalnak viszont volt mit mondania.

Talán a munkastílus segít. Talán elérhető, hogy a könyvtár műveltségkincséhez elvezető irodalomtörténetet hozzáértő szerzőgárda, de mégis szinte kedvtelésből hozza létre. Persze így sem lehet garantálni, hogy az új irodalomtörténeti narratíva vonzó és hű lesz. Akármilyen lesz is, mindenképpen jelentkezik majd más is, új irodalomtörténeti mesével. Ebben áll a könyvtár felelőssége. Nem engedhető meg, hogy a könyvtár mélyébe vezető hiperhivatkozások légióját egyetlen irodalomtörténeti kézikönyv kisajátítsa. A könyvtárban létre kell hozni egy olyan különleges közvetítő felületet irodalomtörténészek számára, amely a könyvtár digitalizált belső anyagára mutató hiperhivatkozásokat tartalmazza, és amely nyitva áll bármely narratívabarkácsoló számára. A könyvtár attól könyvtár, hogy polcain nincs cenzúra. Egymás mellett állnak egymást elutasító gondolatok. A könyvtár minden érkező narratívának szükségképpen megengedi, hogy igénybe vegye hiperhivatkozásai kincstárát. A gondolatok versenyét – azért könyvtár, azért nem szerkesztőség – örökre fenntartja.

Összefoglalásul kimondhatjuk, hogy a korszerű eszközöktől nem idegenkedő filológus tehát igenis kiveheti a részét a digitális örökség megőrzésének – mert szétosztásának – feladatából.

Digitális fordulat

A *digital turn* megérkezett Magyarországra. Egyetemem hírlevelének legutóbbi száma arról értesít, hogy a digitális bölcsészet új, magyar–német kutatócsoportja szép támogatást kapott. A rövid beszámolóban négyszer szerepel a *digitális fordulat*, olykor röviden *fordulat* kifejezés. A digitális bölcsészet talán még a tananyagba is bekerülhet, írják: „a projekt keretében a magyar doktoranduszok aktív részeseivé válhatnak annak, ahogy az ELTE-n megismertetünk és bevezetünk egy olyan kutatási irányt, amely vélhető jövőbeni humántudományos szerepét tekintve egyetemünkön

⁴⁶ „Az új irodalomtörténet »felnőtt« könyv akar lenni: Várkonyi Benedek beszélgetése Veres András szerkesztővel,” *Magyar Tudomány* 167, 4. sz. (2007): 502, <http://www.matud.iif.hu/07apr/14.html>.

egyelőre nincs kellőképpen reprezentálva.”⁴⁷ Az *egyelőre* szót azért furcsállom, mert hiszen a közelmúltig két összefüggő programban is tanítottunk valamit, amit bízvást digitális bölcsészettnek vélhettünk. A két program nem mindenben volt elmaradva a nagyvilágtól. A Digiphil vezetője Aloys Pilcher műhelyének internetes Wittgenstein-forráskiadását állítja elibénk ösztönző példaként⁴⁸ – ám az 1990-es évek végén, amikor a német Wittgenstein-kiadók még sugárlemezen (CD-ROM) publikáltak, a Digiphil mai beosztott kutatói, Mihály Eszter és Parádi Andrea már hálózati kritikai kiadást tettek közzé.⁴⁹ A *Bölcsészettudományi Informatika Önálló Programban* (BIÖP) a hálózati szakkönyv technológiáit fejlesztettük a '90-es évek óta, a reneszánsz doktori és mesterképzési műhelyben pedig 1993 óta az előző információrobbanás, a könyvnyomtatás korát tanulmányozzuk abban a meggyőződésben, hogy a *historia magistra vitae* alapelve mellett ott a *vita magistra historiae* is. A – könyvtáros-továbbképzésnek hála – nyereséges BIÖP működését 2010 körül megszüntették, a reneszánsz mesterszak működésének pedig 2018-ban kell véget vetnünk.

Digitális fordulatról beszélnek – ma, amikor a számítógép alkalmazása immár nem ütközik ellenállásba. De akkor minek nevezzük azt, ami a reneszánszkutatók nyíregyházi konferenciáján, 1980 májusában történt, miután az RPHA első változatát bemutattuk? Figyeljük meg, amit a hazai reneszánsz- és barokk-kutatás jó okkal tisztelt vezetőjének szavaiból a jegyzőkönyv az *Irodalomtörténeti Közlemények* számára rögzített:

Klaniczay Tibor elmondta, hogy ő jóindulatú szkepszissel figyelte a vállalkozást. A számítógép alkalmazásától azt várná, hogy gyorsan kapjon megbízható információt. Csatlakozik Pirnát Antal véleményéhez: csak 100%-ig biztos adatot szabad feldolgozni; e munkában viszont túl sok a hipotetikus adatfelvétel. Tudtával a feldolgozás az RMKT befejező kötetblokkja érdekében készült. Ő a fordított utat tartotta volna helyesebbnek: ha már készen lenne az RMKT XVI. századi sorozata, akkor lehetne jó lelkiismerettel ilyen jellegű munkákba kezdeni. Így viszont a bizonytalan adatokat tároljuk a biztos megállapítása érdekében. Az elmondottak miatt a vállalkozást jobban el tudná képzelni XVII. századi anyagon. Nem hiszi, hogy az előadó ismertette módon irodalomtörténeti problémák megoldhatók lennének. Például arra a már régebben felmerült kérdésre, hogy vajon van-e a XVII. századi szövegek (székelyföldi énekeskönyvek) között középkori, az irodalomtörténész hamarabb megadja a választ, mint a számítógép, ha az adatelőkészítésre fordított időt figyelembe vesszük.⁵⁰

⁴⁷ „Az irodalomtudomány gyakorlata: Digitális fordulat és társadalmi hasznosság,” kutatás ismertető (ELTE BTK, 2016), <http://www.elte.hu/file/Digitalizalodo-irodalomtudomany-DAAD-TKA-ELTE-Koln.pdf>, 7.

⁴⁸ Palkó Gábor, „Mit jelent a digitális filológia.”

⁴⁹ Az idézett Balassi Bálint összes verse, hálózati kritikai kiadás közreműködőiként – lásd a kiadás „Névjegy”-ét. Hozzáférés: 2018.05.15, <http://magyar-irodalom.elte.hu/gepesk/bbom/inevjegy.htm>.

⁵⁰ Kőszeghy Péter és Lázs Sándor, „A magyar irodalom műfajai és a reneszánsz társadalom (Tudományos ülésszak), Nyíregyháza, 1980. május. 14–16.” *Irodalomtörténeti Közlemények* 5–6. sz. (1980): 760–766, http://epa.oszk.hu/00000/00001/00319/pdf/itk00001_1980_05-06_759-771.pdf.

Nem akarom sem szegedi, sem pesti csoportunk érdemeit az egekbe magasztalni. Elmaradásunk az Egyesült Államok kutatóihoz képest elmondhatatlanul nagy. Susan Hockey régi cikkében még ma is elképeszt az a tény, hogy Aquinói Tamás számítógépes szótárának munkálatait Roberto Busa amerikai jezsuita már 1949-ben megkezdte.⁵¹ Csoportunk csak jelentéktelen mértékben tudta gazdagítani a filológiai gyakorlatot, de a rend kedvéért megnevezem igénypontjainkat.

- Az RPHA volt az első sok szempontú, számítógépes költészeti leltár (1992), és elsőként került a világhálóra (1994). Az RPHA adatmodellje eredeti alkotás volt (Gál György, 1988).⁵²
- A valószínűleg első hálózati forráskiadást (*Balassa-kódex*, 1993, ill. 1994) és az első több forrású, hálózati kritikai kiadást (*Balassi Bálint összes verse*, 1998) mi tettük közzé.
- A sztochasztikus olvasás eljárása (véletlen-generátor alkalmazása bizonytalan textológiai helyzetben, 1998).
- Talán annak felismerése is új volt, hogy az emberiség teljes digitális öröksége – kb. 70 kultúrával számolva – tulajdonképpen csekély méretű, így sokmillióss többszörözés révén megőrizhető (2004 és 2005).⁵³

Az új keletű digitális bölcsészet

A *fordulat* szóban, amikor az ember a saját tevékenységére alkalmazza, van valami, ami Attila királyra emlékeztet:

Csillag esik, föld reng: jött éve csudáknak!
Ihol én, ihol én, pörölye világnak!

Marx *A tőkéhez* Dante-mottót választott: „Menj utadon, s ne bánd, hogy mit beszélnek,”⁵⁴ Copernicus viszont, bár „de revolutionibus” írt (az égitestek forgásáról),⁵⁵ nem forradalmárként lépett föl, és műve Galilei koráig nem is került rá a tiltott könyvek pápai jegyzékére.

A *digitális bölcsészet*: ez valami új, különleges bölcsészetet jelent (most zárójelbe teszem, hogy éppen digitálisat) a jelzőtlen bölcsészethez képest. De mi a baj azzal, amit a középkor óta a bölcsészkarokon művelünk? Nem túlzás-e ekkora igénnyel

⁵¹ Susan Hockey, „The History of Humanities Computing,” in *A Companion to Digital Humanities*, eds. Susan Schreibman, Ray Siemens and John Unsworth (Oxford: Blackwell, 2004), <https://doi.org/10.1111/b.9781405103213.2004.x>.

⁵² Gál György, „A »Répertoire de la poésie hongroise ancienne« adatmodellje,” *Irodalomtörténeti Közlemények* 93, 3. sz. (1989): 267–272.

⁵³ 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-book*ot (f = 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.

⁵⁴ Karl Marx, *A tőke: A politikai gazdaságtan bírálata*, ford. Rudas László és Nagy Tamás (Budapest: Szikra Könyvkiadó, 1955), 15.

⁵⁵ Nicolaus Copernicus, *De revolutionibus orbium coelestium* (Nürnberg: Johannes Petreius, 1543).

fellépni? Az *Association for Computers in the Humanities* 1978-as megalapításakor még megelégedett azzal, hogy a bölcsészeti kutatásokat számítógéppel segítse, de nem akart új, másfajta bölcsészetet létrehozni. Ma már ez a szervezet is, mint mindenki, a *digital humanities* kifejezést használja. A BIÖP megalapításakor nem taktikai okból kerültük a forradalmi beszédmódot. Nem mindenki akar fordulatot a bölcsészetben. 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 bölcsészetnek viszont lehetőleg ne kelljen alkalmazkodnia semmihez, így a számítógép mindenkori lehetőségeihez sem. Amikor – a számítógép kora előtt más tudományokban is szokatlan – teljes indukciós bizonyítással vezettük be az *izo-szabályt*, előre láttuk, hogy a teljes indukcióra való új képességünk magával fogja hozni a – mai szóval – *távoli olvasás* szélsőséges fajtáit is, egészen az anyagismeretről való lemondás, a nem-olvasás filológiai lehetőségéig. Nem voltunk elragadtatva ettől a jövőtől, de nem is tartottuk elkerülhetőnek. Példával mutattuk be, hogy az, aki az RPHA felépítését nem jól ismeri, szükségképpen eltéved benne. Az 1536-os esztendő csak amiatt látszik verstörténetünk oly kiemelkedő évének, mert a számítógép bosszút áll az anyagát nem eléggé ismerő filológuson, és a szemébe hazudik. És nemcsak mi nem bántottuk a bölcsészetet, hanem ő sem bántott minket. A mindenkori főnökök más kérdés, de a kollégáktól mindig csak segítséget: elismerést vagy jogos bírálatot kaptunk.

Áttanulmányozva a digitális bölcsészet körül zajló amerikai vitát, úgy látom, hogy ez nem csupán reklám, hanem valódi fordulat, amely a bölcsészet lényegét illeti. Eddig a bölcsészet könyvek körül zajlott, és volt egy ókori eredetű, bevált módszerünk annak megállapítására, hogy mi található az illető könyvekben. Az új, digitális bölcsészet is könyvek körül zajlik, de új módszerrel tesz megállapításokat a könyvekről.

Tömegével vizsgálják a műveket, mint az orosz formalisták előfutárunktól, Ivan Veszelszkiától Vlagyimir Proppig – vagy mint mi az RPHA-adatbázis segítségével az 1601 előtti magyar verseket. Lehet, hogy nagyképűség, de érteni vélem az irányzat megalapítóját, Franco Morettit, sőt közel érzem magamhoz. Mintha szerepelne egykori könyvemben, *A vers*⁵⁶ harmadik fejezetében. Módszerét, a *távoli olvasást* szembeállítja a francia és német metafizika ihlette irodalomelmélettel. Amivel közösséget vállal, az a természettudomány, a kemény társadalomtudományok és az ezzel szerinte jól összeegyeztethető marxizmus.⁵⁷ Én a Lotz Jánostól tanult *T-oppozíciót* használtam. A metafizikus irodalomelméleteket (első fejezet) magam is szembeállítottam a tudománnyal (második fejezet), bár a marxizmust nem a tudományba tettem, hanem beolvasztottam a metafizikák közé. Tudományos és metafizikus irodalomszemlélet oppozícióját – tehát magát a szembenállást – opponáltattam az eloszlásokat vizsgáló, nem-olvasó, esetleg még a szóban forgó irodalom nyelvét sem ismerő irodalomtörténésszel (harmadik fejezet), aki nem szokta ilyen fennkölt dolgokon törni a fejét.

A szembeállítások rendszere tehát nem azonos, de rokon, és az is rokon, hogy *A versben* e megközelítéseket egymást kiegészítőknak, szükségesnek, sőt ciklikusan, felváltva alkalmazandóknak mutattam be. 2016. márciusi interjújában Moretti elmeséli,

⁵⁶ Horváth Iván, *A vers*, 2000 Könyvek (Budapest: Gondolat Kiadó, 1991), 188–192.

⁵⁷ Franco Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History* (London: Verso, 2005), 2.

hogy rávette kiadóját arra, hogy két utolsó könyvét⁵⁸ ugyanaznap jelentesse meg, mert a két könyv – a digitális bölcsészet módszerét alkalmazó *Distant Reading* és az ettől a módszertől nagyon távoli *The Bourgeois* „ugyanannak az éremnek két oldala.” A két megközelítés kizárja egymást („mutually exclusive”), de az irodalomtörténésznek az általuk feltárt mindkét valóság egyformán fontos.⁵⁹

Moretti vonakodik attól, hogy meghatározza a digitális bölcsészetet („digital humanities means nothing”), a kibic azonban az amerikai vita⁶⁰ nyomán a következő alapsémát véli látni. Beláthatatlan digitális adattömeg jött és még jön létre a könyvtárakban, és a digitális bölcsész ezt az adattömeget nem arra a célra használja, amire létrehozták, hanem mindenféle rendezéseket figyel meg rajta. A nagy különbség az RPHA létrehozói és a digitális bölcsészek között az, hogy – bár mi is csináltunk ún. *távoli olvasás* típusú vizsgálatokat, amelyekhez a műveket el sem kellett olvasnunk – az anyagot mi magunk, többek között épp ilyen vizsgálatok számára hoztuk létre. A digitális bölcsész készen találja anyagát, és kihasználja azt. (Vegyünk egy tényleges kutatást. A 19–20. század fordulóján számos klasszikus szerző kapott írói szótárt. A kutató megvizsgálja, hogy annak a klasszikusnak a szakirodalma, aki szótárt kapott, gyorsabban kezdett-e el növekedni az új segédeszköz birtokában, mint annak a klasszikusnak a szakirodalma, aki nem kapott. A válasz – amelyet a kutató meglepőnek minősít – negatív.)⁶¹

A könyvtárban a régi vágású digitális filológus rendezéseket és narratívákat (szakbibliográfiákat, hálózati kritikai kiadásokat, sokszempontú leltárakat, irodalomtörténeteket) hoz létre, ha úgy tetszik: maga is építi a virtuális könyvtárat – a fordulat utáni digitális bölcsész pedig vad, önálló rendezéseket valósít meg a mások által létrehozott anyagon, és így olyasmit is hozzáférhetővé tesz a bölcsészeti kutatás számára, amiről ezt a lehetőséget korábban nem sejtettük, vagyis új használatra mutat példát. A virtuális könyvtárnak elvileg mindkét bölcsészeti tevékenységből haszna lehet – az elsőből persze több. Ha valakinek nem tetszik a digitális bölcsész, mert olyasmiről beszél, amihez nem ért,⁶² ha élődinek tartja, ha irigykedik rá, az abban a körülményben találjon magának vigaszt, hogy az interjú szerint az előkelő stanfordi

⁵⁸ Franco Moretti ed., *Distant Reading* (London: Verso, 2013); Franco Moretti, *The Bourgeois: Between History and Literature* (London: Verso, 2013).

⁵⁹ Az idézetek forrása: Melissa Dinsman, „The Digital in the Humanities: An Interview with Franco Moretti,” *Los Angeles Review of Books*, 2016. márc. 2., <https://lareviewofbooks.org/interview/the-digital-in-the-humanities-an-interview-with-franco-moretti/>.

⁶⁰ Matthew K. Gold, ed., *Debates in the Digital Humanities* (University of Minnesota Press, 2012). Európában a kifejezés jelentése talán elmosódottabb. Csak a szövegkódolás elméletével foglalkozik: Jacques Dubucs, „Digital Humanities: Foundations,” in *Exploring the Paradigm Shift*, ed. Péter Davidházi (Amsterdam: Amsterdam University Press, 2014), 21–35. A darmstadti Műegyetem digitális bölcsészet-programja az „új eszközök és módszerek” híve: <http://www.digitalhumanities.tu-darmstadt.de/index.php?id=35&L=1>.

⁶¹ Holger S. Syme and Scott Selisker, „In Defense of Data: Responses to Stephen Marche’s »Literature Is not Data«,” *Los Angeles Review of Books*, 2012. nov. 5., <https://lareviewofbooks.org/essay/in-defense-of-data-responses-to-stephen-marches-literature-is-not-data>. A szerző, Scott Selisker többi példája is hasonló.

⁶² Stephen Marche, „Literature Is not Data: Against Digital Humanities,” *Los Angeles Review of Books*, 2012. nov. 5., <https://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities>.

professzornak épp úgy nincs programozója, neki is épp úgy a hallgatója programoz, ahogy nekem Pesten.

Digital Humanities in the Virtual National Library

A virtual national library can be distributed among people in arbitrary number of copies. Duplication is the only way of long-term preservation since the last millenials have shown that the durability of text carriers seems to have decreased as the number of copies have grown. Another condition of long-term preservation is that people also use the national cultural heritage to which they have access. This can be facilitated by providing narratives (such as literary history), not solely texts. “Digital philology” follows the classification principles of traditional philology when registering, editing the works and contextualizing them in a narrative. The first milestones of Hungarian digital philology: the repertory of the early modern Hungarian poetry: 1994, the first online critical edition: 1998, the first online Hungarian literary history: 2006. In recent years digital humanities scholars have highlighted hitherto unknown correlations and patterns in literary history thanks to these novel digital editions as well. These correlations are sometimes relevant, sometimes not. My book *A vers* (*The Verse*, 1991) foresaw the possibility of a “non-reading” philologist who uses Franco Moretti’s distant reading method.

Keywords:

virtual library, national library, digital philology

Lejtovicz Katalin

Österreichische Akademie der Wissenschaften (ÖAW), Austrian Centre for Digital Humanities (ACDH)

Katalin.Lejtovicz@oeaw.ac.at

Bernád Ágoston Zénó

ÖAW, Institut für Neuzeit- und Zeitgeschichtsforschung (INZ)

Agoston.Bernad@oeaw.ac.at

Matthias Schlögl

ÖAW, ACDH

Matthias.Schloegl@oeaw.ac.at

Maximilian Kaiser

ÖAW, INZ

Maximilian.Kaiser@oeaw.ac.at

Peter Alexander Rumpolt

ÖAW, Institut für Stadt- und Regionalforschung (ISR)

Peter.Rumpolt@oeaw.ac.at

Digitalizáció és hálózatkutatás: Az Österreichisches Biographisches Lexikon 1815–1950 és az APIS-projekt*

Az elmúlt években fejlődésnek induló digitalizálási módszerek lehetővé tették a digitális bölcsészet területén belül többek között az életrajzi lexikonokat feldolgozó projektek előtérbe kerülését. A biográfiai kutatások célja a már digitalizált lexikonok szövegeinek (pl. *Österreichisches Biographisches Lexikon 1815–1950* – a továbbiakban *ÖBL* – részben strukturált, részben strukturálatlan, szöveges formában tárolt adatainak) átalakítása számítógép által feldolgozható formátumra, illetve az átalakított adatokon hálózati, statisztikai és számítógépes nyelvészeti kutatások elvégzése. Ebben a kontextusban kezdődtek el 2015-ben a kutatások az Osztrák Tudományos Akadémia Új- és Jelenkortörténeti Kutatási Intézetében [Institut für Neuzeit- und Zeitgeschichtsforschung] a Mapping historical networks: Building the new Austrian Prosopographical, Biographical Information

* A tanulmány az Osztrák Tudományos Akadémia Új- és Jelenkortörténeti Kutatási Intézetében, az osztrák Nationalstiftung für Forschung, Technologie und Entwicklung által támogatott Mapping historical networks: Building the new Austrian Prosopographical, Biographical Information System (APIS) című projekt keretében készült.

System (APIS) projekt keretén belül. A projekt az ÖBL digitális feldolgozását és egy kutatási célokra kifejlesztett számítógépes környezet létrehozását célozza meg, mely a szintén az Osztrák Tudományos Akadémia keretében működő Austrian Centre for Digital Humanities (ACDH) [Osztrák Digitális Bölcsészeti Központ] valamint az Institut für Stadt- und Regionalforschung (ISR) [Város- és Régiókutatási Intézet] együttműködésével valósul meg. Cikkünkben bemutatjuk, hogy az APIS-projekt technikai megvalósítására fejlesztett webalkalmazás milyen technikai és nyelvtechnológiai megoldásokat használ a lexikon számítógépes feldolgozására és az adatok reprezentálására.

Kulcsszavak:

életrajzi lexikon, információ kinyerés, entitás kapcsolat, történeti hálózatkutatás



1. Bevezetés

Az Osztrák Tudományos Akadémia Új- és Jelenkortörténeti Kutatási Intézetében 2015 óta zajlanak a Mapping historical networks: Building the new Austrian Prosopographical, Biographical Information System (APIS) című projekt keretében végzett kutatások. A projekt célja az *Österreichisches Biographisches Lexikon 1815–1950* (ÖBL)¹ című életrajzi lexikon új, digitális rendszerben való feldolgozása, mely a szintén az Osztrák Tudományos Akadémia keretében működő Osztrák Digitális Bölcsészeti Központ valamint a Város- és Régiókutatási Intézet együttműködésével valósul meg. A modern információtechnológiai környezetnek és a számítógépes nyelvészeti adatfeldolgozásnak köszönhetően az ÖBL egyrészt a szemantikus web Giant Global Graph csomópontjává válhat,² ugyanakkor pedig a kulturális ismeretanyag tárolójaként is funkcionálhat, amely új kérdésfelvetéseket és megoldási javaslatokat tehet lehetővé a szellem-, társadalom- és kultúratudományok területén, különösen pedig a migráció- és a történeti hálózatkutatások körében. Jelen tanulmány elsőként az ÖBL (digitalizálási) történetét vázolja fel, majd bemutatja az APIS-adatbázis jelenlegi változatát, és rávilágít a számítógépes nyelvészeti eljárásokra, végezetül pedig betekintést enged néhány történeti jellegű kutatásba is.

2. Az ÖBL története és digitalizálási fázisai

Az ÖBL kezdetei a 20. század elejéig nyúlnak vissza. 1916-ban kezdeményezte az irodalomtörténész Anton Bettelheim (1851–1930) és a történész Oswald Redlich (1858–

¹ *Österreichisches Biographisches Lexikon 1815–1950*. Lfg. 68, Bd. 14 (Wien: Böhlau Verlag, Verlag der Österreichischen Akademie der Wissenschaften, 1954ff.). *Österreichisches Biographisches Lexikon 1815–1950 Online Edition*, 2003–2017, hozzáférés: 2017.12.14, <http://www.biographien.ac.at/>.

² Tim Berners-Lee, James Hendler and Ora Lassila, „The Semantic Web: A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities,” *Scientific American*, 2001. máj. 17., 34–43; Dave Raggett, „W3C Data Activity: Building the Web of Data,” hozzáférés: 2017.12.14, <http://www.w3.org/2013/data/>.

1944) a Constant von Wurzbach által szerkesztett, 1856 és 1891 között megjelenő *Biographisches Lexikon des Kaiserthums Österreich*³ című életrajzi lexikon folytatását és átdolgozását. A tervet a történelmi helyzet miatt csak részben lehetett megvalósítani,⁴ az összegyűjtött forrásanyag viszont lehetővé tette, hogy a második világháborút követően elkezdődhessen egy újabb életrajzi szakmunka összeállítása: az Osztrák Állami Levéltár igazgatója, Leo Santifaller (1890–1974) sürgetésének köszönhetően 1946-ban az Osztrák Tudományos Akadémia határozatot hozott az *ÖBL* kidolgozására. A lexikon első füzeté 1954-ben jelent meg.

A lexikonba azon személyek adatai kerülnek be, akik 1815 és 1950 között az egykori osztrák államszerkezetben éltek vagy ott hunytak el, és a közélet valamelyik területén jelentőset alkottak. Ezáltal az *ÖBL* a határokon átívelő nemzeti biográfiai munkák paradigmatis példája lesz, hiszen adatait nem szűkíti a mai Ausztriára, hanem összefogja az egykori császári birodalom területeit, illetve az Osztrák–Magyar Monarchiában tevékenykedők adatait. A lexikon ily módon a középeurópai kultúra leképezőjeként is funkcionál, mely átfogóan mutathatja be e régió kultúrájának 19. és 20. századi változásait. A nyomtatott változat jelenleg 14 kötetből áll (68 füzet), A és V betűk között több mint 18 000 életrajzot közöl. A mű magyar vonatkozásaival kapcsolatban kiemelendő, hogy a kötetek több mint 2500 olyan személy adatait tartalmazzák, akik az egykori Magyar Királyság vagy Erdély területén születtek. Ezáltal az *ÖBL* a magyar és erdélyi történelmi kutatások egyik legfontosabb német nyelvű biográfiai forrásmunkájaként konstituálódik.

A lexikon, illetve a szerkesztőség által gyűjtött nyilvántartólapok a digitalizálás több fázisán vannak túl. Az 1990-es években az addig kartotékokon rögzített adatokat bevezették egy Microsoft Access adatbázisba. Ezt a 2000-es évek elején egy OracleTM alapú relációs adatbázis váltotta fel, amely 450 000 rekordot tartalmaz, amelyből több mint 80 000 életrajzi adatnak tekinthető.⁵ Ennek a továbbra is rendelkezésre álló, platformfüggetlen integrációt megvalósító digitális rendszernek köszönhetően jött létre az *ÖBL* online kiadása 2009-ben azáltal, hogy mind a strukturált, mind pedig a nem strukturált személyi adatok átvételét és XML-formátumba való átmentését lehetővé tette. Az XML-formátum rögzített formában tartalmazza a nevet, születési és elhalálozási adatokat (a születési és elhalálozási helyet is), foglalkozást, műveket

³ Constant von Wurzbach, Hg., *Biographisches Lexikon des Kaiserthums Oesterreich, enthaltend die Lebensskizzen der denkwürdigen Personen, welche seit 1750 in den österreichischen Kronländern geboren wurden oder darin gelebt und gewirkt haben*, Bd. I–LX (Wien: L. C. Zamarski, k. k. Hof- und Staatsdruckerei, 1856–1891).

⁴ Anton Bettelheim, Begr., *Neue Österreichische Biographie 1815–1918, Abteilung 1: Biographien*, Bd. I–VIII (Wien–Zürich: Wiener Drucke–Amalthea Verlag, 1923–1935); Anton Bettelheim, Begr., *Neue Österreichische Biographie ab 1815: Große Österreicher*, Bd. IX–XXII (Wien–München: Amalthea Verlag, 1957–1987); Hans Bohatta, Jaroslav Sutnar und Julius Stockinger, Hg., *Neue Österreichische Biographie 1815–1918, Abteilung 2: Bibliographie zur Neuen Österreichischen Biographie*, Bd. I (Zürich–Wien–Leipzig: Amalthea Verlag, 1925).

⁵ Christine Gruber und Roland Feigl, „Von der Karteikarte zum biografischen Informationsmanagementsystem: Neue Wege am Institut Österreichisches Biographisches Lexikon und biographische Dokumentation,” in *Biografische Lexika im Internet: Internationale Tagung der „Sächsischen Biografie” in Dresden (30. und 31. Mai 2008)*, Hg. Martina Schattkowsky und Frank Metasch, Bausteine aus dem Institut für Sächsische Geschichte und Volkskunde 14 (Dresden: Thelem Universitätsverlag, 2009), 55–75.

és vonatkozó irodalmat. A legtöbb információ azonban az életrajzi szócikkek nem strukturált folyó szövegében található.

3. Az APIS-projekt technikai aspektusai

Az új digitalizálási lehetőségeknek és módszereknek köszönhetően az elmúlt években az életrajzi lexikonokat feldolgozó projektek előtérbe kerültek. Ezeket a munkálatokat egy közös cél vezérli: a már digitalizált biográfiai kézikönyvek szövegeinek (melyek az adatokat félig strukturált formában tárolják) átalakítása a gépek által feldolgozható adatokká (pl. Deutsche Biographie,⁶ BiographyNet),⁷ illetve a hálózati kutatások módszereinek alkalmazása ezen adatok értelmezésében.⁸ Ebben a kontextusban kezdődtek el az ÖBL-ben található biográfiák számítógépes nyelvészeti feldolgozásai is az Institut für Maschinelle Sprachverarbeitung der Universität Stuttgart (Stuttgarteri Egyetem Gépi Nyelvfeldolgozás Intézet) falai között.⁹ A következőkben bemutatjuk, hogy az APIS-projekt¹⁰ technikai megvalósítására fejlesztett webalkalmazás milyen technikai és nyelvtechnológiai megoldásokat használ a lexikon számítógépes feldolgozására és az adatok reprezentálására.

3.1. Az adatmodell

Az ÖBL-ben található adatok leírására egy saját adatmodellt fejlesztettünk ki. Már létező modellek felhasználását, mint például a CIDOC-CRM¹¹ elvetettük, mert ezek jellemzően generikusak, a digitális bölcsészet témakörébe tartozó sokféle adat leírására lettek kifejlesztve. Ha egy ilyen, már létező, generikus adatmodellt használtunk volna az ÖBL adatainak jellemzésére, túl komplex reprezentációhoz jutunk, ahol az olyan egyszerű események, mint például születés, házasságkötés, elhalálozás többszörösen elágazó struktúrákat eredményez (legalább négy csomópont CIDOC-CRM-ben). Az

⁶ Deutsche Biographie, hozzáférés: 2017.12.14, <https://www.deutsche-biographie.de/>.

⁷ BiographyNet, hozzáférés: 2017.12.14, <http://www.biographynet.nl/>.

⁸ Sophia Stotz, Valentina Stuß, Matthias Reinert and Maximilian Schrott, „Interpersonal Relations in Biographical Dictionaries: A Case Study,” in *BD2015: Biographical Data in a Digital World 2015, CEUR Workshop Proceedings*, vol. 1399, eds. Serge ter Braake, Antske Fokkens, Ronald Sluijter, Thierry Declerck and Eveline Wandl-Vogt (Amsterdam: CEUR-WS.org, 2015), 74–80, <http://ceur-ws.org/Vol-1399/paper12.pdf>; Antske Fokkens, Serge ter Braake, Niels Ockeloen, Piek Vossen, Susan Legêne, Guus Schreiber and Victor de Boer, „BiographyNet: Extracting Relations Between People and Events,” in *Europa baut auf Biographien: Aspekte, Bausteine, Normen und Standards für eine europäische Biographik*, Hg., Ágoston Zénó Bernád, Christine Gruber und Maximilian Kaiser (Wien: New Academic Press, 2018), 193–224.

⁹ André Blessing, Andrea Glaser and Jonas Kuhn, „Biographical Data Exploration As a Test-bed for a Multi-view, Multi-method Approach in the Digital Humanities,” in *BD2015*, 53–60, <http://ceur-ws.org/Vol-1399/paper9.pdf>; Jonas Kuhn und André Blessing, „Die Exploration biographischer Textsammlungen mit computerlinguistischen Werkzeugen: Methodische Überlegungen zur Übertragung komplexer Analyseketten in den Digital Humanities,” in *Europa baut auf Biographien*, 225–257.

¹⁰ APIS (GitHub), hozzáférés: 2018.01.09, <https://github.com/acdh-oeaw/apis-core>.

¹¹ CIDOC-CRM: Conceptual Reference Model, hozzáférés: 2018.01.09, <http://www.cidoc-crm.org/>.

ehhez hasonló, komplex modellek megnehezítik egy webalkalmazás fejlesztését, karbantartását és bővítését. A mi esetünkben elegendő az osztályoknak, az osztályokat jellemző tulajdonságoknak és az osztályokat összekötő relációknak egy részhalmaza az életrajzok leírására. Az APIS-projektben öt entitást definiáltunk: Személy, Hely, Intézmény, Esemény, Mű. Ezek az osztályok tovább bonthatók, tipizálhatók a felhasználó által definiált ontológiákkal, továbbá az osztályok közötti kapcsolatok típusa is leírható hierarchikus kontrollált szótárakkal.

Művek esetében például a felhasználó megadhatja az adott mű típusát: irodalmi, képzőművészeti vagy tudományos stb. mű, ezeken belül az ontológia által definiált további kategóriákat rendelhet az entitásokhoz, például: festmény, szobor vagy fénykép stb. A tipizálás pontosságát a definiált szótár leíróképessége határozza meg. Az entitások mellett a relációk is elláthatóak típussal, melyek szintén a felhasználó által definiált struktúrából választhatóak ki. A személy–személy osztályok közötti kapcsolatok elláthatóak *házastárs*, *szülő*, *tanár* stb. típussal, a személy–intézmény kapcsolatok a *tanult*, *dolgozott* stb. típusokkal. Az APIS-projektben az osztályok közötti kapcsolatok leírására használt típusok mindig kétirányúak, ezt tükrözik a relációk elnevezései is: *parent of*, *child of* („szülő”, „gyerek”), *works in*, *has employee* („dolgozik valahol”, „munkavállaló”).

3.2. Az APIS-webalkalmazás

Az APIS-alkalmazás¹² segítséget nyújt a kutatóknak nagy mennyiségű adathalmaz szisztematikus és félautomatikus feldolgozásában, az adatok elemzésében, és az ennek során megtalált összefüggések megjelenítésében. Az adatok megjelenítése lehetővé teszi, hogy az egyébként strukturálatlan adatokban nehezen átlátható összefüggések felszínre kerüljenek, továbbá, hogy a kutatók könnyebben és intuitívabban dolgozhassanak nagy mennyiségű adatokkal a felhasználóbarát megjelenítésnek köszönhetően.

Az APIS-projekt keretében egy széles körben alkalmazható VRE-t (Virtual Research Environment, Virtuális Kutatói Környezet) fejlesztettünk ki, amely az indulás óta számos kutatási területen és több intézetben is felhasználásra került. Az APIS egy *Django*¹³ és *MySQL*¹⁴ frameworkre épülő webalkalmazás, amely nemcsak felhasználóbarát felhasználói felülettel, hanem egy RESTful API-val is rendelkezik, lehetővé téve, hogy a felhasználók ne csak a felhasználói felületen keresztül, hanem programkódból is hozzáférhessenek az adatokhoz. Az adatmodell által leírt információ a webalkalmazásban kerül megjelenítésre. Az alkalmazás lehetővé teszi bejegyzések egyszerű létrehozását, módosítását és törlését. A felület három fő komponensből áll: az első, az adatok különböző kritériumok szerint szűrhető listája, a második az entitások tulajdonságait megjelenítő nézet, a harmadik pedig a vizualizációs és fájlletöltés funkciók weboldala. A felhasználók az adatokat a részletes leírást tartalmazó nézetben tudják szerkeszteni; az oldalt úgy terveztük meg, hogy a felhasználót

¹² Matthias Schlögl und Katalin Lejtovicz, „APIS – Eine Linked Open Data basierte Datamining-Webapplikation für das Auswerten biographischer Daten,” in *DHd 2017 Bern: Digitale Nachhaltigkeit: Konferenzabstracts*, eds. Michael Stolz et al. (Bern: Universität Bern, 2017), 220–222.

¹³ Django Software Foundation, *Django*, hozzáférés: 2018.01.04, <https://www.djangoproject.com/>.

¹⁴ *MySQL*, hozzáférés: 2018.01.04, <https://www.mysql.com/>.

támogassa az adatok szerkesztésében. A bal oldalon található a kiválasztott bejegyzés metaadatai (pl. születési hely, idő), jobb oldalon pedig az aktuális entitás összeköthető más bejegyzésekkel (pl. egy személy összeköthető az intézménnyel, ahol tanult).

Teleki von Szék, László d. J. Gf.

Texts

Metadata Person

Name

Teleki von Szék

Start

11.2.1811

Please enter a date (DD).(MM).YYYY

End

7.5.1861

Please enter a date (DD).(MM).YYYY

Status

First name

László d. J. Gf.

The persons's forename. In case of more then one name...

Profession

Literatur, Buch- und Zeitungswesen

Politik

Politiker und Schriftsteller

Title

Type to get suggestions

Gender

male

MetaInfo

related to

Event

Institution

Delete	Start Date	End Date	relation type	Related Institution	Edit
<input type="checkbox"/>	01.01.1836	01.01.1844	war Mitglied >> war k. Mitglied	Ungarische Akademie der Wissenschaften	Edit
<input type="checkbox"/>	01.01.1844	—	war Mitglied >> war Ehrenmitglied	Ungarische Akademie der Wissenschaften	Edit
<input type="checkbox"/>	—	—	war Mitarbeiter >> war Praktikant	Ungarische Statthaltere	Edit
<input type="checkbox"/>	01.01.1828	01.01.1830	war in Ausbildung >> war Student	Reformiertes Kollegium Sárospatak	Edit
<input type="checkbox"/>	—	—	war Mitarbeiter >> war Praktikant	Ungarische Hofkanzlei	Edit
<input type="checkbox"/>	—	—	war Mitglied >> war Vizepräsident	Ungarischer Schutzverein (Védegylet)	Edit
<input type="checkbox"/>	—	—	war Mitglied >> war Präsident	Pester Kreis (Pesti Kör)	Edit
<input type="checkbox"/>	—	—	war Mitglied >> war Präsident	Oppositionskreis (Ellenzéki Kör)	Edit
<input type="checkbox"/>	—	—	war Mitglied >> war Abgeordneter	Ungarischer Landtag	Edit
<input type="checkbox"/>	—	—	war Gründer	Radikalkreis (Radical Kör)	Edit

Weiter

Seite 1 von 2

Relation type*

Type to get suggestions

Institution*

Type to get suggestions

1. ábra. Az életrajzban bemutatott személy metaadatait és kapcsolatait megjelenítő webes felület az APIS-alkalmazásban

Az ilyen entitás–entitás kapcsolatok egyszerű létrehozását a webes felület *auto-complete* (automatikus kiegészítés) funkcióval támogatja, továbbá a földrajzi helyek (falvak, városok, országok) egy térképen jelennek meg, ezzel segítve a felhasználót a megfelelő hely kiválasztásában.

Az APIS-alkalmazásban nemcsak webes felületet lehet használni az adatok bevitelére, hanem lehetőség van annotációk megadására a szabad szövegben is. A felhasználó az egér segítségével kijelöli az annotálni kívánt szövegrészt, majd a felugró contextmenüből kiválasztja a reláció típusát (pl. *utazott valahova*). A kiválasztott reláció mindig az életrajzban szereplő személyt köti össze egy másik entitással. A reláció kiválasztásakor a másik entitás típusának megfelelő webes felület automatikusan felugrik, a felhasználó ebben adja meg – automatikus kiegészítés segítségével – a kijelölt szövegrészhez tartozó adatbázis bejegyzést (ebben az esetben pl. egy város vagy ország nevét). A megjelölt reláció az adatbázisban elmentésre kerül, továbbá a szövegben

szinkóddal emeljük ki az annotációt a könnyebb kereshetőség és módosíthatóság érdekében.

Default
All selected (17)
All selected (20)
Update

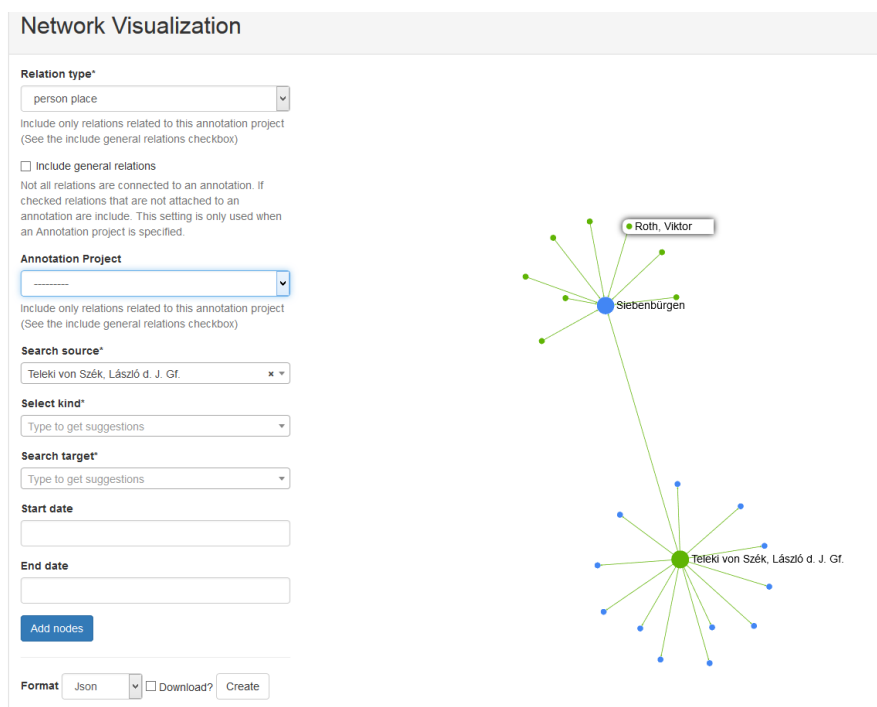
Sohn von → László Gf. T. v. S. d. Á., Halbbruder von → József Gf. T. v. S. d. J. – Nach privatem Unterricht absolv. T. 1828–30 Jus am ref. Kollegium von Sárospatak. Darauf folgte ein kurzes Praktikum bei der Statthaltere in Ofen (Buda), dann in Wien bei der Ung. Hofkanzlei. Stud.reisen führten ihn 1833–36 nach Dtd., in die Niederlande, nach England und Frankreich. Nach seiner Rückkehr wandte sich T. der Politik zu und wurde eines der aktivsten Mitgl. der liberalen Reformopposition in Siebenbürgen und Ungarn. Sein literar. Erstlingswerk war das Drama „Kegyenc“, 1841. In den 1840er-Jahren war T. eine der führenden und gleichzeitig radikalsten Figuren der hochadeligen Opposition. 1844 wurde er zum Vizepräs. des Schutzver., 1845 zum Präs. des Pester Kreises gewählt, der ersten Organisation der liberalen Opposition. 1847 wurde er Präs. des Oppositionskreises (Ellenzéki Kör). Auf dem letzten ständ. LT verlangte T. an der Seite von → Ludwig Gf. Batthyány die sofortige Inkraftsetzung des Reformprogramms. Während der Revolution schuf er aus dem Oppositionskreis den „Radikalkreis“ und wurde in Abony als Abg. in den LT gewählt. Im Repräsentantenhaus galt er als geistiger Anführer einer Opposition, die die Batthyány-Regierung von links kritisierte und entschiedenes Auftreten gegenüber Wien forderte. Obwohl T. als ung. Gesandter in Paris für sein Land keine diplomat. Anerkennung erreichen konnte, blieb er dennoch dort, um die ung. Interessen in Westeuropa weiterhin entscheidend zu vertreten. T. verstand es, die ausländ. Medien geschickt für die Zwecke der ung. Revolution einzusetzen. Nach Besprechungen mit den Anführern der poln. und rumän. Emigration entwarf er den Plan für ein föderatives Ungarn als Zentrum einer neuen Donau-Konföderation, worin die Nationalitätenkonflikte gelöst werden sollten. Nach der Niederlage der Revolution hielt sich T. ab 1851 größtenteils in der Schweiz auf. Während des Krieges in Italien gründete er Anfang Mai 1859 gem. mit → Lajos Kossuth v. Udvard u. Kossut und → Georg Klapka die Emigrantenregierung „Ungarisches Nationaldirektorium“. Im November 1860 reiste er – aus privaten Gründen – illegal nach Dresden. Dabei wurde er von der sächs. Polizei verhaftet und an Osterr. ausgeliefert. Nach zehntägiger Haft wurde er unter der Bedingung freigelassen, sich von polit. Tätigkeiten fern zu halten. Doch schon im April 1861 nahm er erneut als Abg. von Abony am parlamentar. Leben teil, wobei er jedes Zugeständnis gegenüber Osterr. zurückwies und erklärte, die Zustände von 1848 als Ausgangspunkt für weitere Verh. betrachten zu wollen. Im Unterschied zur „Adresspartei“ → Franz v. Deáks forderte T., dass dieser Vorschlag dem Kg. per Beschluss – daher „Beschlusspartei“ – überbracht würde. T. wurde Anführer der Letzteren, doch schreckten die Mitgl. vor notwendigen Zugeständnissen gegenüber den nichtmagyar. Nationalitäten zurück und so blieb T. mit seiner vergleichsweise konzilianten Haltung innerhalb der Partei allein. Ab 1836 war er K. M., ab 1844 Ehrenmitgl. der MTA.

Teleki von Szék László Gf. d. J., Politiker und Schriftsteller. Geb. Pest (Budapest, H), 11. 2. 1811; gest. ebd., 7. (8.) 5. 1861 (Selbstmord); evang.

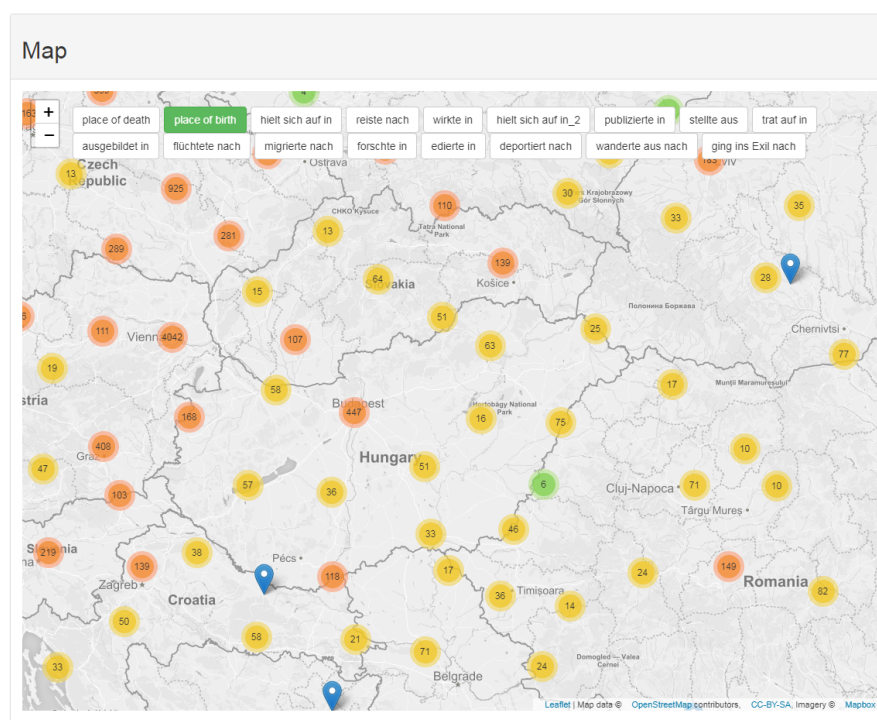
2. ábra. Az életrajzok annotációját támogató webes felület az APIS-alkalmazásban

Az APIS-webalkalmazás fejlesztésekor nemcsak arra törekedtünk, hogy az adatok beírása és szerkesztése egyszerű legyen, hanem hogy a nagy mennyiségű adathalmazban kódolt információt a kutatók gyorsan és könnyen megtalálják, és egy átfogó képet kapjanak róla. Ennek érdekében több vizualizáció is a felhasználó rendelkezésére áll. Az egyik az entitások közötti relációkat bemutató hálózat, melyben a személyek, helyek, intézmények, művek, események a hálózat csomópontjai, ezeket kötik össze a relációkat reprezentáló élek, melyeken minden esetben szerepel a reláció típusa. A gráf személyre szabható, lehetőség van a megjelenítendő csomópontok és élek típusának kiválasztására. A kutatók ezáltal megjeleníthetik a számukra releváns információkat, például az összes személyt, aki egy bizonyos intézményben dolgozott. Ez lehetővé teszi, hogy egy átfogó képet kapjanak az adatokról, anélkül, hogy át kellene tekinteniük egy nagyméretű szövegtörzset. Továbbá az így megjelenített adatokat exportálni lehet, és le lehet tölteni JSON, valamint GraphML fájlformátumokban, így lehetőség nyílik az információ megosztására, továbbá nyílt forráskódú vizualizációs szoftverek használatára.

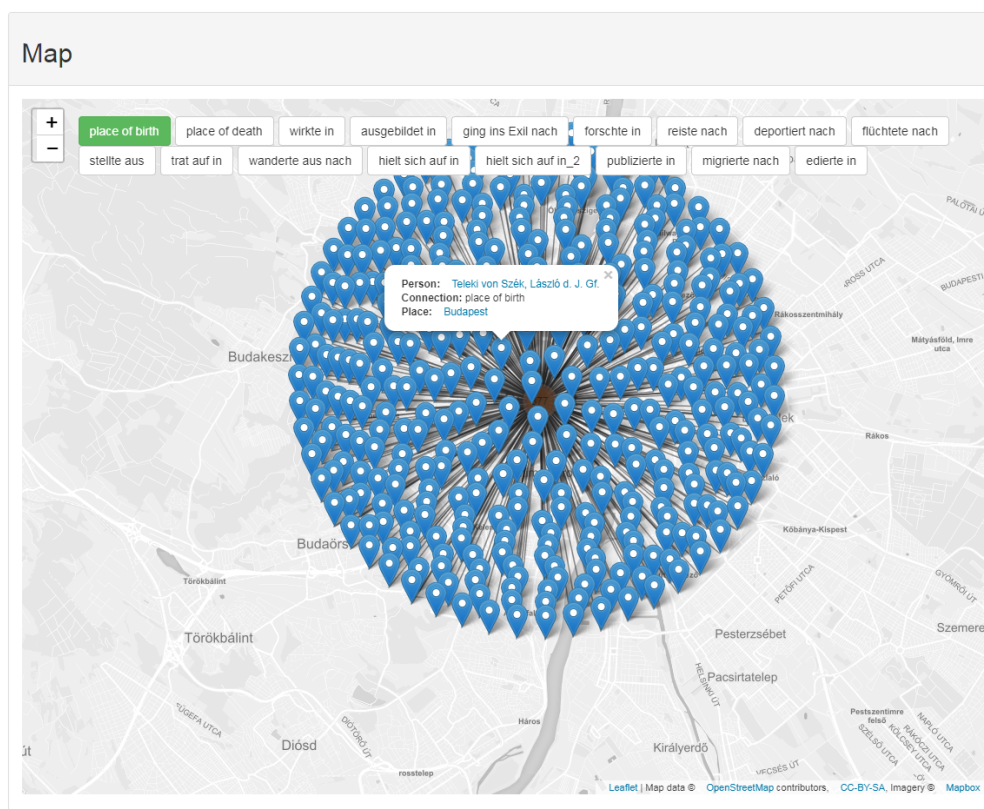
Egy másik vizualizációs lehetőség az APIS-ban az életrajzokhoz köthető helyek térképen történő megjelenítése. A térkép nézete aszerint változtatható, hogy a felhasználó milyen helyek megjelenítését választja ki a rendelkezésre álló típusok közül (pl. születési hely, tartózkodási hely stb.). Az egyes helységekre ráközelítve a felhasználó az áttekintést nyújtó nézetből (amelyen körökbe írt számok mutatják, hogy hány életrajz kapcsolódik egy adott városhoz vagy területhez) átválthat a személyeket egyesével bemutató nézetre.



3. ábra. A személy–hely kapcsolatokat bemutató hálózati ábra ifjabb gróf Teleki László (1811–1861) életrajza alapján



4. ábra. Az ÖBL-be felvett személyek születési helyei – térképvizualizáció az APIS-adatbázisban (részlet)



5. ábra. Budapesten született személyek az APIS-adatbázis térképén

3.3. Információkinyerés

Az APIS-projekt egyik célja, hogy a kutatók munkáját megkönnyítő automatikus feldolgozási lehetőségeket kínáljon, amelyek biztosítják, hogy az életrajzok szövegében található információ automatikusan és hatékonyan a felhasználó rendelkezésére álljon. A szövegek feldolgozását és értelmezését számítógépes nyelvészeti módszerekkel végeztük el, amelyek magukban foglalják a biográfiákban előforduló entitások (személyek, helyek, intézmények stb.) azonosítását, ezeknek az automatikus hozzákapcsolását a Linked Open Data Cloud ¹⁵ adatforrásaihoz, a kapott URI-k automatikus egyértelműsítését, az entitások közötti relációk felismerését és az eredmények manuális kiértékelését. A következő pontokban ezeket a lépéseket fogjuk részletesen bemutatni.

3.4. Névelem-felismerés és entitás-összekapcsolás

Bár az ÖBL-ben az életrajzok XML-formátumban is rendelkezésre állnak, ezek a fájlok néhány kulcsfontosságú eseményt (pl. születés és elhalálozás) kivéve nem tartalmaznak strukturált adatokat a leírt személyek életéről. A projekt egyik fő célja, hogy a természetes nyelvű szövegben kódolt információkat, például a személyek, helyek, in-

¹⁵ Linked Data – Connect Distributed Data Across the Web, hozzáférés: 2018.01.04, <http://linkeddata.org/>.

tézmények, események stb. neveit és az életrajzban bemutatott személynek az ezekhez fűződő viszonyát automatikusan feltárja. Ahhoz, hogy ezt hatékonyan és mindemellett jó minőségben végezhessük el, automatikus és a manuális információkinyerési technikákat ötvöztünk.

Az APIS-ban egy olyan információ kinyerésére alkalmas automatikus rendszert hoztunk létre, amely alkalmazható más digitális bölcsészeti projektek természetes nyelvű szövegeinek annotálására. Az általunk használt rendszer a következő komponensekből áll: a névellem-felismerő (Named Entity Recognizer), az entitás-összekötő (Entity Linker) és az egyértelműsítő komponens. Az automatikus információkinyerésre az Apache Stanbol¹⁶ nevű nyílt forráskódú szoftvert használtuk, amely a természetes nyelvű szövegben megtalált entitásokat kapcsolja össze a referenciaként szolgáló ontológiákkal és tudástárakkal, mint például a GND (Gemeinsame Normdatei),¹⁷ a GeoNames¹⁸ vagy a DBpedia.¹⁹ Az összeköttetések, amelyek az entitások és az életrajzok között így létrejönnek, nemcsak a szövegek szemantikai információkkal történő kiegészítését, hanem a hiányzó vagy hibás adatok kijavítását, illetve pótlását is lehetővé teszik.

A projektben használt Apache Stanbol előnye, hogy bármilyen RDF/XML-formátumban rendelkezésre álló ontológiát fel lehet használni a szövegek szemantikus kiegészítésére, továbbá biztosítja a dokumentumok NLP (Natural Language Processing, természetesnyelv-feldolgozás) szoftverekkel történő feldolgozását. A szemantikus annotálás elvégzéséhez az RDF/XML-formátumban rendelkezésre álló adatokból úgynevezett Referenced Site-okat²⁰ hozunk létre, ezek egy Solr-indexben²¹ tárolják az ontológiák adatait, ami *cache*-elten vagy a szerveren lokálisan érhető el.

Az APIS-ban létrehozott folyamat, amely a strukturálatlan, szöveges életrajzokból strukturált, szemantikai információkkal kiegészített és számítógépesen feldolgozott dokumentumokat hoz létre, két lépésből áll. Először feloldjuk a rövidítéseket, köztük a személynevek, intézménynevek, az akadémiai címek, a helynevek és a gyakori igék rövidítéseit. Két változatát is kifejlesztettük a rövidítések feloldásának. Az egyik, egy reguláris kifejezéseken alapuló Java program, ami az ÖBL-ben használt rövidítések listáját használja. A másik megoldás egy Python szkript, amely reguláris kifejezéseket és egy nagy méretű német nyelvű korpuszt használ a többjelentésű rövidítések feloldására és a helyes megoldás kiválasztására. A program lekérdezi a rövidítéseket és azok környezetét (előtte és utána álló szavakat) a német nyelvű adathalmazban, és amelyik változatra a legtöbb találatot kapja, azt a rövidítésfeloldást választja.

A szemantikus annotálási folyamat második lépése a Solr-indexek létrehozása az ontológiákból. Az entitás-összekapcsolás (Entity Linking) során ezekben az indexekben keres az Apache Stanbol a szövegben megtalált entitásoknak (személyek, helyek, intézmény nevei stb.) megfelelő fogalmakat az ontológiákban. Az APIS-projektben

¹⁶ Apache Stanbol, hozzáférés: 2018.01.04, <https://stanbol.apache.org/index.html>.

¹⁷ Deutsche Nationalbibliothek, *Gemeinsame Normdatei*, hozzáférés: 2018.01.04, http://www.dnb.de/EN/Standardisierung/GND/gnd_node.html.

¹⁸ GeoNames, hozzáférés: 2018.01.04, <http://www.geonames.org/>.

¹⁹ DBpedia, hozzáférés: 2018.01.04, <http://wiki.dbpedia.org/>.

²⁰ „Working with Custom Vocabularies,” hozzáférés: 2018.01.04, <https://stanbol.apache.org/docs/trunk/customvocabulary.html>.

²¹ Apache Solr, hozzáférés: 2018.01.04, <http://lucene.apache.org/solr/>.

a GeoNames-ből és a GND-ből készítettünk indexeket, hogy a szövegben előforduló helyneveket, személyneveket és intézmények neveit hozzá tudjuk kapcsolni a Linked Open Data Cloud-hoz. Az indexeket a következőképpen hoztuk létre: letöltöttük a fent említett források RDF/XML-formátumú *dump*jait, amelyeket egy saját fejlesztésű szkripttel felváltunk részfájlokkra. A fájlok feldarabolásának két oka, hogy kezelhető méretű bemenetet kapjunk, és hogy könnyen tudjunk külön indexeket készíteni a különböző típusú entitásoknak. Az így kapott kisebb méretű fájlokból az Apache Stanbol indexelésre létrehozott Java package-ével készítettük el az Apache Solr-indexeket és az OSGi-keretrendszerben²² futó OSGi *bundle*-t (OSGi-köteget).

A Solr-index és az OSGi-köteg létrehozása és installálása után történik az Entity Linking komponens konfigurálása, amely a következő, kötelezően végrehajtandó lépésekből áll: az Entity Linking komponens azonosítására egy egyedi név megadása, a használni kívánt forrás (pl. GeoNames vagy GND) és a forrás Solr-indexében a kereséshez használni kívánt mező kiválasztása. Ezeken felül további beállítások megadására is van lehetőség, ami pontosabb és hatékonyabb entitás-összekapcsolást (*entity linkinget*) tesz lehetővé. Szűkíthetjük például csak a tulajdonnevekre a keresést, ekkor az algoritmus csak a tulajdonneveket választja ki a szövegből, és ezeket kérdezi le a Solr-indexben, így optimalizálva a találati precizitást és a futási időt. Lehetőség van továbbá entitástípusok leképezésére is, amennyiben az index tartalmaz erre vonatkozó információt. A leképezés eredményeként a felhasználó az indexben különböző típusokkal ellátott entitásokat átfogó kategóriákba sorolva kapja meg az alkalmazástól, így például összevonva, a *Személy* kategória alatt lehet megmutatni az összes, az indexben férfiként, illetve nőként megjelölt nevet.

Az Entity Linking komponens konfigurációját követően a bemeneti adatok feldolgozására vonatkozó nyelvtechnológiai információkat adjuk meg. Az alkalmazás számára ez írja le, hogy milyen NLP-lépéseket végezzen el, és azokat milyen sorrendben hajtsa végre a természetes nyelvű szövegeken. Az APIS-ban az Apache OpenNLP²³ nyílt forráskódú szoftvert használjuk az életrajzok számítógépes nyelvészeti feldolgozására. A szövegfeldolgozás folyamata a következő lépésekből áll:

- A bemeneti szöveg nyelvének megállapítása.
- A szöveg mondatokra tagolása (opennlp-sentence).
- A mondatok szavakra bontása (opennlp-token).
- A szavak szófajának megállapítása (opennlp-pos).
- A főnévi kifejezések megkeresése (opennlp-chunker).
- Entitás-összekapcsolás. (Entity Linking.)

Utolsó lépésként az NLP-feldolgozás során kapott főnevek és főnévi kifejezések a Solr-indexszel kerülnek összehasonlításra (*entity linking*). Amennyiben egy főnévi kifejezés megegyezik az indexben tárolt entitások valamelyikével, az entitást és a hozzá tartozó információkat az alkalmazás a kért formátumban szerializálva adja vissza. Ha több találat is van, 0-tól 1-ig terjedő pontszám jelzi, hogy melyik a legvalószínűbb helyes eredmény. Az általunk használt Entity Linking-szoftver előnye,

²² OSGi Alliance, hozzáférés: 2018.01.04, <https://www.osgi.org/>.

²³ Apache OpenNLP, hozzáférés: 2018.01.04, <https://opennlp.apache.org/>.

hogy tetszőleges RDF/XML-formátumban rendelkezésre álló ontológiát hatékonyan tud indexálni, így a felhasználó kiválaszthatja, hogy az alkalmazás melyik adatforrással annotálja meg a bemeneti szövegtörzset.

3.5. Relációfelismerés

Az entitás-összekapcsolás az első lépés a természetes nyelvű szövegek jelentésének automatikus értelmezésében. Segítségével a szövegben szereplő személyneveket, helyneveket, intézmények neveit össze tudjuk kapcsolni a Linked Open Data Cloud szótáraival. Ennek az előnye, hogy a szövegben karaktersorozatokként szereplő neveket egyedi azonosítóval (Uniform Resource Identifier, URI) rendelkező entitások váltják fel. Az ontológiák URI-val rendelkező fogalmai nemcsak egyértelműen azonosíthatók és hivatkozhatók, megoszthatók a különböző alkalmazások között, hanem használatukkal lehetőség nyílik a strukturálatlan szövegek információval történő kiegészítésére, illetve a helytelen adatok kijavítására.

A második lépés a korpuszok szemantikus feldolgozásában a megtalált entitások közötti kapcsolatok és azok típusainak megállapítása, más néven az automatikus relációfelismerés.

A relációfelismerés során az NLP-modul a szövegben detektált személyek, helyek, események között keres szemantikai kapcsolatokat, mint például *szülő-gyerek*, *utazott valahova*, *tanult valahol*, *részt vett egy eseményen* kapcsolat. Az automatikus relációfelismerésre több módszert is kipróbáltunk, melyek közül tesztelést követően fogjuk kiválasztani az alkalmazásba véglegesen integrált megoldást. Az első változat egy szabályalapú algoritmus, melyet a GATE-keretrendszer²⁴ használva implementáltunk. Az implementáció a GATE-annotációkon is működő reguláris kifejezési nyelvet, a JAPE-et²⁵ használja a szemantikai kapcsolatok szövegből történő automatikus kivonatolására. Első lépésként az Entity Linking modul kimenetét XML-formátumba konvertáltuk, ahol minden entitás külön elemként jelenik meg, az elemek attribútumai pedig az entitást tartalmazó Solr-index neve (pl. GeoNames_Capital_Cities_Index, GND_Persons_Index) és az entitáshoz köthető (pl. GeoNames vagy GND) URI.

Az XML-fájlokat ezután betöltöttük GATE-be, majd a beépített NLP-modulokkal tokenizáltuk és azonosítottuk a szövegekben a névelemeket. Az importált fájlok entitásai (az entity linking kimenete) és a GATE-ben végrehajtott NLP-feldolgozás eredménye (tokenek, névelemek) annotációkként tárolódnak GATE-ben. A JAPE reguláris kifejezések az annotációkon végeznek műveleteket, a különböző relációkra jellemző mintázatokat keresve a szövegekben. Amennyiben az alkalmazás talál egy, az adott relációra jellemző mintázatnak megfelelő szövegrészt, azt automatikusan ellátja egy újabb annotációval, amely a reláció típusát írja le. A relációazonosítás kimenetét XML-fájlokba – az NLP alkalmazások által széles körben használt formátumba – exportáljuk. A második relációfelismerési megoldás, amelyet teszteltünk, a nyílt for-

²⁴ GATE: General Architecture for Text Engineering, hozzáférés: 2018.01.04, <https://gate.ac.uk/>.

²⁵ „JAPE: Regular Expressions over Annotations,” hozzáférés: 2018.01.04, <https://gate.ac.uk/sale/tao/splitch8.html>.

ráskódú, Pythonban implementált *IEPY*-szoftver,²⁶ amely gépi tanuláson alapuló relációfelismerést végez. Az alkalmazás webes felületén a felhasználó annotálja az általa definiált relációkat (pl. *utazott valahova, megházasodott valakivel*), majd a szoftver az így létrehozott korpuszból megtanul egy modellt, amely az entitások között fennálló relációk felismerésére alkalmas. Az *ÖBL* esetében az *IEPY* nem bizonyult megfelelő szoftvernek, mert a relációk mindkét tagjának bejelölését megköveteli (pl. a *tanult valahol* esetében a személy- és helynevet egyaránt), az *ÖBL*-ben azonban az ismétlések elkerülése érdekében sokszor csak egyszer, az életrajz elején szerepel a személynév.

A harmadik megközelítés, amelyet megvizsgáltunk, a mondatok szintaktikai elemzésével (*parszolásával*) kapott fastruktúrából mélytanulás (Deep Learning) módszerrel történő relációfelismerés,²⁷ ahol a bemenetet néhány száz annotált biográfia adja. Ennek a megoldásnak a hátránya, hogy nagyon nagy mennyiségű annotált adatot kíván meg a tanulási folyamat, amely jelenleg nem áll rendelkezésünkre. A későbbiekben a reguláris kifejezésen és a Deep Learning módszeren alapuló megoldásokat fogjuk összehasonlítani egy erre a célra létrehozott tesztkorpuszon, hogy a pontos eredményességi mutatók alapján (precision, recall, F1-score) döntsük el, melyik megoldást választjuk.

3.6. Egyértelműsítés és manuális szöveggondozás

A fent bemutatott automatikus szövegfeldolgozási módszerek jellemzője, hogy bár könnyen és hatékonyan képesek természetes nyelvű szövegeket feldolgozni, az eredmények sok esetben helytelenek. A hibás eredmények két formája az APIS-ban, amikor egy entitást vagy relációt nem talál meg az alkalmazás (pl. nem ismer fel a szövegben egy személynevet), illetve amikor a megtalált entitás vagy reláció azonosítása helytelen (pl. *tanult valahol* helyett a *dolgozott valahol* relációt adja vissza a program). A szoftver pontosságának növelése érdekében a két típushibát egyrészt automatikus, másrészt manuális módszerekkel próbáljuk kiküszöbölni. A helytelenül azonosított városok, országok és egyéb helynevek automatikus javítására a következő heurisztikákat alkalmazzuk:

Amennyiben több találat is érkezik az Entity Linking modultól, azt választjuk, amelyik:

- magasabb adminisztratív egységnek felel meg, azaz egy fővárost előnyben részesítünk például egy kisvárossal, faluval szemben. (Pl. a *Bécsben járt egyetemre* mondat esetében *Wient*, Ausztria fővárosát választjuk az USA-ban található *Wien* kisvárossal szemben.)
- A nagyobb lélekszámú település (azonos adminisztratív szintű helységek esetében nagyobb a valószínűsége, hogy a népesebb a helyes megoldás.)
- Közelebb van az *ÖBL*-ben releváns földrajzi területhez, az Osztrák-Magyar Monarchiához (nagyobb valószínűséggel éltek, dolgoztak az életrajzokban szereplő

²⁶ *IEPY: Information Extraction Tool*, hozzáférés: 2018.01.04, <https://github.com/machinalis/iepy>.

²⁷ Matthias Schlögl, Katalin Lejtovicz, Ágoston Zénó Bernád, Maximilian Kaiser and Peter Rumpolt, „Using Deep Learning to Explore Movement of People in a Large Corpus of Biographies,” *Zenodo*, 2018. jan. 16., <http://doi.org/10.5281/zenodo.1149023>.

Záluží, kr. Středočeský

Texts

Metadata Place

Name

Záluží, kr. Středočeský

Start

Please enter a date (DD).(MM).YYYY

End

Please enter a date (DD).(MM).YYYY

Status

ambiguous

Kind

Latitude

Longitude

Metadata

delete

modify

Revisions

Linked Open Data

related to

Map

Search for...

Search

+

-

Place

ÖBL name: Záluží, kr. Středočeský

Geonames: Záluží

Select this URI

Confidence: 37.65585

Feature: <http://sws.geonames.org/3061899/>

Prague

CHKO Český kras

Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

Use sequences of search terms (delimiter: .) to limit the search to certain administrative divisions. E.g. "Österreich, Oberösterreich, Steyr". Use square brackets to set the min score difference to slice results. "Wien" will only show the capital of Austria, "Wien[30]" shows also a town in India.

Institution

Person

Event

Place

Work

6. ábra. Az egyértelműsítést támogató felhasználói felület az APIS-ban. A többértelmű helységnevek térképen jelennek meg a döntést segítő plusz információkkal (név, GeoNames URI)

emberek a Monarchia területén, vagy az ahhoz közeli országokban, mint távolabbi területeken.)

A fenti heurisztikák sem tudják minden esetben a helyes megoldást kiválasztani, ezért az APIS-webalkalmazás lehetővé teszi a találatok kézi javítását, és a többértelmű találatok egyértelműsítését. Az alábbi ábra az egyértelműsítésre kifejlesztett felhasználói felületet mutatja be.

Az automatikus és kézi hibajavítás ötvözése az APIS-ban megfelelő megoldásnak bizonyult az adatok pontos és hatékony feldolgozására, de hogy minél gyorsabban és minél kevesebb emberi erőforrással tudjunk nagy mennyiségű szöveges adatot feldolgozni, a jövőben tervezzük a kézi hibajavítás automatizálható részeinek implementálását.

4. Kutatási témák

Az APIS-projekt egyik célkitűzése az ÖBL-ben publikált biográfiai adatok alapján végzett történeti népesség- és migrációkutatás. A több mint 18000 életrajzot tartalmazó

korpusz szellem- és társadalomtudományi szempontból történő vizsgálatát – beleértve a humángéográfia tárgykörébe tartozó, főként népességföldrajzi és mobilitás-, illetve migrációkutatási kérdésfeltevéseket – egyrészt egy transzdiszciplináris kutatócsoport együttműködése, másrészt a digitalizált életrajzi szócikkekre épülő virtuális kutatási környezet (Virtual Research Environment, VRE) teszi lehetővé.

A lexikonba felvett személyek migrációs útvonalainak és mintáinak rekonstrukciójára összpontosító szociodemográfiai elemzés egyúttal egy készülőben lévő disszertációs dolgozat témája. Az *ÖBL* életrajzi szócikkei ugyanis térvonatkozású információkat is tartalmaznak (születési és halálozási hely, tartózkodási helyek stb.), melyek segítségével rekonstruálhatók a térbeli életutak. Ebben a kontextusban mind a térbeli mobilitás mérete, illetve a migráció, a migrációs útvonalak és vándorlási minták, mind pedig az érintett népességcsoportok demográfiai struktúrája figyelmet érdemel, és a kutatás tárgyát képezi. Emellett vizsgáljuk a városok mint migrációs csomópontok szerepét, valamint a települések és az intézmények hatókörét. Az életrajzokban említett, de nem egyértelműen identifikálható helyneveket történeti kartográfiai források segítségével azonosítjuk. Az elemzés összességében a 19. század elejétől a 20. század közepéig tartó korszak politikai és politikai-adminisztratív változásainak keretei között helyezhető el. Ebben az összefüggésben értékelhetők a projekt eddig publikált részeredményei is, melyek az egyes személyek migrációjának példáján keresztül a földrajzi nevek közötti összefüggésekre világítanak rá a közösségkeresési módszerek (*community detection*) segítségével.²⁸

Az életrajzi szócikkekben előforduló helynevek jelentőségét a migrációs folyamatok vizsgálata szempontjából idősebb gróf Andrássy Gyula (1823–1890) és ifjabb gróf Teleki László (1811–1861) életrajza alapján szemléltetjük.²⁹ Andrássy 1867–1871 között magyar miniszterelnök, 1871–1879-ben pedig a Monarchia külügyminisztere volt. Az *ÖBL* Andrássy-szócikke számtalan térvonatkozású adatot tartalmaz: többek között Kassát mint születési helyet, Pestet, illetve a pesti egyetemet mint tanulmányai színhelyét, Isztambult mint egyik munkahelyét, Londont és Párizst, ahol száműzetésben élt, valamint az isztriai Voloskót, ahol elhunyt.

²⁸ Ágoston Zénó Bernád, Maximilian Kaiser, Sebastian M. Mair und Alexander Rind, „Communities in biographischen Netzwerken,” in *FMT 2017: Proceedings of the 10th Forum Media Technology and 3rd All Around Symposium, CEUR Workshop Proceedings*, vol. 2009, eds. Wolfgang Aigner, Thomas Moser, Kerstin Blumenstein, Matthias Zeppelzauer, Michael Iber, Grischa Schmiedl (St. Pölten: CEUR-WS.org, 2017), 83–87, <http://ceur-ws.org/Vol-2009/fmt-proceedings-2017-paper12.pdf>; Ágoston Zénó Bernád, Maximilian Kaiser, Matthias Schlögl and Katalin Lejtovicz, „APIS Dataset Humanities Scholars,” *Zenodo*, 2017. szept. 19., <https://doi.org/10.5281/zenodo.896281>.

²⁹ „Andrássy Julius (d. Ältere) Graf,” in *Österreichisches Biographisches Lexikon 1815–1950*, Bd. I., Hg. Leo Santifaller und Eva Obermayer-Marnach (Graz/Köln: H. Böhlau, 1957), 20–21; Zoltán Fónagy, „Teleki von Szék László Gf. d. J.,” *Österreichisches Biographisches Lexikon 1815–1950*, Bd. 14, Hg. Ernst Bruckmüller und Christine Gruber (Wien: Verlag der Österreichischen Akademie der Wissenschaften, 2015), 135–136.

Andrássy, Julius d. Ä. Gf.

Texts

Metadata Person

Name

Andrássy

Start

8.3.1823

Please enter a date (DD).(MM).YYYY

End

18.2.1890

Please enter a date (DD).(MM).YYYY

Status

First name

Julius d. Ä. Gf.

The persons's forename. In case of more then one name...

Profession

Politik

Ministerpräsident

Title

Type to get suggestions

Gender

male

related to

Event

Institution

Person

Place

Delete	Start Date	End Date	relation type	Related Place	Edit
<input type="checkbox"/>	—	—	place of birth	Košice	<input type="button" value="Edit"/>
<input type="checkbox"/>	—	—	place of death	Volosko	<input type="button" value="Edit"/>
<input type="checkbox"/>	—	—	ausgebildet in	Pest	<input type="button" value="Edit"/>
<input type="checkbox"/>	01.01.1847	01.01.1847	wirkte in	Pressburg	<input type="button" value="Edit"/>
<input type="checkbox"/>	01.01.1848	01.01.1848	wirkte in	Schwechat	<input type="button" value="Edit"/>
<input type="checkbox"/>	01.01.1849	01.01.1849	wirkte in	Istanbul	<input type="button" value="Edit"/>
<input type="checkbox"/>	01.01.1849	—	reiste nach >> ging ins Exil nach	London	<input type="button" value="Edit"/>
<input type="checkbox"/>	—	01.01.1857	reiste nach >> ging ins Exil nach	Paris	<input type="button" value="Edit"/>
<input type="checkbox"/>	08.06.1867	08.06.1867	wirkte in	Buda	<input type="button" value="Edit"/>

Relation type*

ausgebildet in

Place*

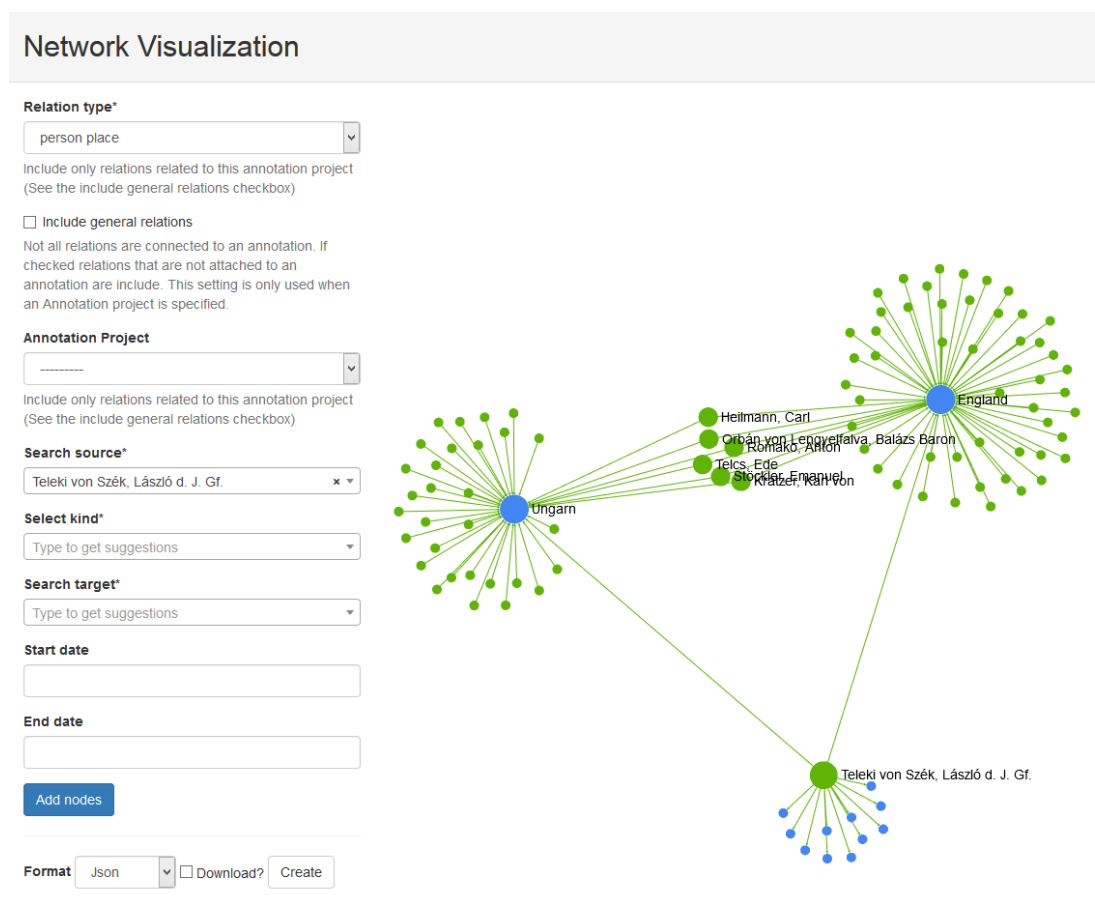
GND Pest (37.823387): Stadt links der Donau, 1873 mit Buda und Óbuda zu Budapest...

7. ábra. Andrássy személy–hely kapcsolatait megjelenítő webes felület az APIS-alkalmazásban

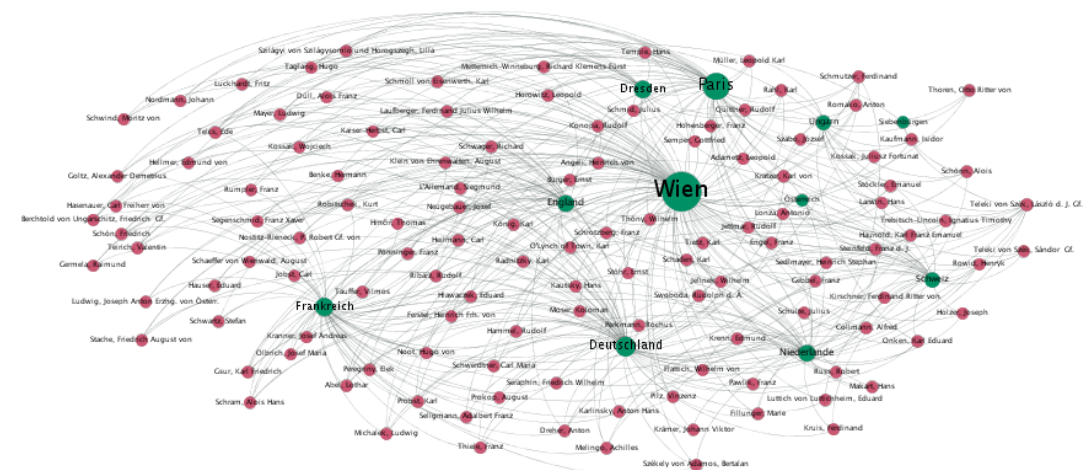
Teleki, az 1848–49-es forradalom és szabadságharc egyik vezető alakja, a Határozati Párt politikusa, a Magyar Tudományos Akadémia tiszteleti tagja volt. Pesten született és halt meg, Sárospatakon, Budán és Bécsben tanult. Ezt követően európai tanulmányútja során bejárta Németországot, Hollandiát, Angliát és Franciaországot. Politikai pályafutásával kapcsolatosan a szócikkben említett földrajzi nevek Erdély, Magyarország, Svájc és Ausztria, illetve Abony, Párizs és Drezda.

További kutatások témáját képezik a művészek és gyűjtők életrajzai, sajtóhálózatok feltérképezése és más tudománytörténeti kérdésfelvetések. A művészi életrajzok esetében mintegy 506 személyből, a bécsi Künstlerhaus tagjaiból álló csoportról van szó, az összetartozásukat éppen ez a tagság biztosítja.³⁰ Két fő szempontot érvényesítünk a kutatáskor: egyrészt különös figyelmet fordítunk a művészek életútjának és karriertörténetének összehasonlító elemzésére, másrészt a lexikonban található, ám ehhez a csoporthoz nem tartozó további személyeket is műgyűjtőként határozhatunk meg. Ez utóbbi esetben egy még feltárandó csoportról van szó, amelynek tagjai között

³⁰ Maximilian Kaiser, „Was uns Biographien über Künstlernetzwerke sagen. Konzepte für eine historische Netzwerkanalyse auf Basis biographischer Texte aus dem *Österreichischen Biographischen Lexikon (ÖBL)*,” in *Europa baut auf Biographien*, 383–403.



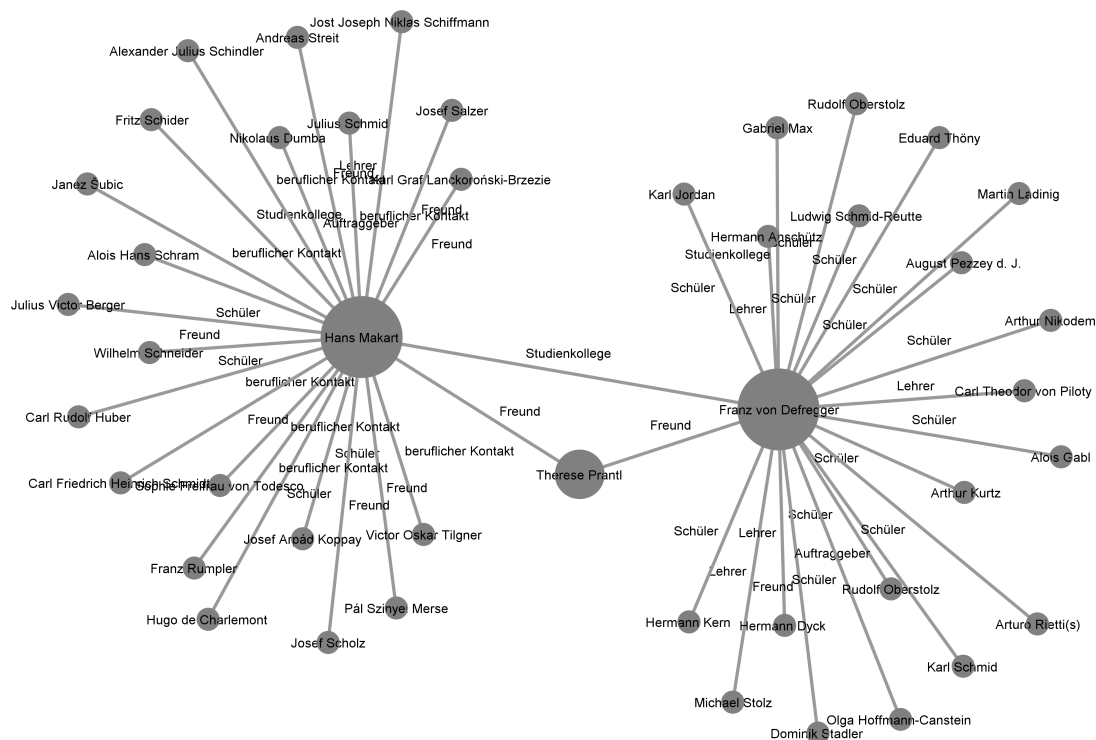
8. ábra. A Teleki-életrajzban előforduló földrajzi nevek hálójának vizualizációja az APIS-adatbázisban (részlet)



9. ábra. A Teleki-életrajzban előforduló földrajzi nevek és a hozzájuk kapcsolódó személyek hálójának megjelenítése a *Gephi* segítségével³¹

³¹ *Gephi: The Open Graph Viz Platform*, hozzáférés: 2017.12.14, <https://gephi.org/>.

orvosokat, ügyvédeket, kereskedőket és nagyiparosokat találunk. Ők megbízóként vagy a műalkotások felvásárlójaként nem elhanyagolható mértékben befolyásolták az osztrák művészettörténet fejlődési folyamatait.



10. ábra. Hans Makart (1840–1884) és Franz von Defregger (1835–1921) festőművészek – a *Visone*-alkalmazás segítségével³² megjelenített – kapcsolati hálója, az *ÖBL*-életrajzokban említett személynevek alapján (© Maximilian Kaiser)

Választásunk azért esett a bécsi Künstlerhaus közösségének vizsgálatára, mert ez jelentette a Habsburg Monarchia legfontosabb reprezentációs és műkereskedelmi platformját a 19. századi művészek számára. A Künstlerhaus kiállításain való részvétel egy fiatal művész számára nem csupán a tagfelvétel és az azzal összefüggésben álló integrálódás miatt volt döntő jelentőségű, hanem a további anyagi sikert is meghatározta. Sok esetben ez biztosította számukra a művészetekre fogékony nyilvánosság előtti megjelenést, és így kerültek kapcsolatba a műgyűjtőkkel is. Az intézménynek kiemelt szerepe van a művészi biográfiák szempontjából igen jelentős műgyűjtői életrajzok esetében is, melyeknek forrásai a Künstlerhaus tagjegyzékei. A csoportok közötti hasonlóságok az életrajzi adatok összehasonlításából és értékeléséből derülhetnek ki. A szövegekben gyakran előforduló helyek, intézmények és személyek nevei olyan hálózatokat rajzolnak ki, amelyek a csoportok közötti kapcsolatokat is kontúrozni tudják.

A 19. és 20. század sajtóhálózatainak vizsgálata a tágabb értelemben vett közép-európai szellem- és kultúrtörténeti kutatások, valamint a történeti hálózatkutatás

³² *Visone: Visual Social Networks*, hozzáférés: 2017.12.14, <http://visone.info/>.

kereteibe illeszkedik. Ebben az esetben nem csupán a kézi annotálással ellátott *ÖBL*-beli életrajzokat, valamint az *ÖBL*-hez tartozó GIDEON-adatbázis kiegészítő adatait használjuk, hanem külső sajtótörténeti forrásokat is bevonunk a kutatásba. Nem törekszünk az intézmények és szereplők összefonódásainak és kapcsolatainak teljes körű feltárására, hanem inkább egy reprezentatív minta feldolgozása a célunk. A kutatás a személyközi viszonyok mellett a személy–intézmény és az intézmények közötti relációkra koncentrál, és azokra a hasonlóságokra és különbségekre kérdez rá, amelyek a Monarchia két államának területén a sajtó és a politikai szféra kölcsönös függőségét és ezen kapcsolat megjelenési formáit jellemezte. Az életrajzok feldolgozásánál feltűnt, hogy a kenyérkereseti foglalkozástól függetlenül igen sok személy jelentős sajtótevékenységet is kifejtett, ennek következtében a továbbiakban fokozottan figyelünk az egyes foglalkozási csoportok közötti többszörös átfedésekre.

Projektünk egyik további aspektusát képezi az *ÖBL* szerzői hálózatának feltérképezése. Ebben az esetben a lexikon létrejöttének körülményei kerülnek előtérbe, ami az *ÖBL* adatainak kritikai feldolgozásához és értelmezéséhez járul hozzá. Mindemellett további adatok bevonása és levéltári kutatások is szükségesek ahhoz, hogy feldolgozhassuk és rekonstruálhassuk azokat a tanár–diák, illetve a személyek és intézmények közötti kapcsolatokat, valamint azoknak a különböző kül- és belföldi szerzői nemzedékeknek a tudományos, politikai és ideológiai szocializációját, amelyek az *ÖBL* általános narratíváját meghatározták. Ezen a téren első kutatási eredményeink időben a hidegháború szakaszára koncentrálnak, és ezen periódusban az *ÖBL*, valamint a magyarországi,³³ illetve romániai³⁴ tudományos intézmények közötti együttműködéseket vizsgálják.

Digitisation and Network Research: The Austrian Biographical Dictionary and the APIS project

Digitisation methods have significantly evolved in recent years and made it possible for all types of digital humanities projects, among them biographical projects to gain ground. One of the aims of a biographical research is to transform the texts of already digitized encyclopaedias (e.g. the *Österreichisches Biographisches Lexikon 1815–1950*, *ÖBL*, which consists of structured and unstructured data) to a machine readable format, and to perform network-, statistical- and computational linguistic analysis on the transformed data. The biographical research of *ÖBL* started in 2015 in the Institute for Modern and Contemporary Historical Research [Institut für Neuzeit- und Zeitgeschichtsforschung (INZ)] of the Austrian Academy of Sciences within the project Mapping historical networks:

³³ Ágoston Zénó Bernád, „Europäische Biographik während des Kalten Krieges. Zur Geschichte der Kooperation zwischen dem Österreichischen Biographischen Lexikon und wissenschaftlichen Institutionen in Ungarn 1956–1989,” in *Europa baut auf Biographien*, 81–105.

³⁴ Ágoston Zénó Bernád, „Das Österreichische Biographische Lexikon 1815–1950. Ein Nachschlagewerk im Wandel der Zeit,” in *Hortus amicorum. Köszöntőkötet Egyed Emese tiszteletére*, szerk. Bartha Katalin Ágnes, Biró Annamária, Demeter Zsuzsa és Tar Gabriella-Nóra (Kolozsvár: Erdélyi Múzeum-Egyesület, 2017), 457–467.

Building the new Austrian Prosopographical, Biographical Information System (APIS). The project is aimed at the digital processing of the *ÖBL* and the creation of a virtual research environment which is implemented in cooperation with the Austrian Center for Digital Humanities (ACDH) and the Institute for Urban and Regional Research [Institut für Stadt- und Regionalforschung (ISR)]. This paper discusses the technical and natural language processing solutions used in the APIS web application to open up biographical data for the researchers by processing and representing them in a structured format.

Keywords:

entity linking, biographical lexicon, information retrieval, historical network research

Cséve Anna*Petőfi Irodalmi Múzeum*

cseve.anna@pim.hu

Fellegi Zsófia*Petőfi Irodalmi Múzeum*

fellegi.zsofia@pim.hu

Kómár Éva*Magyar Nemzeti Múzeum*

komar.eva@mnmu.hu

Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadása: Esettanulmány

2016-ban indult el a Petőfi Irodalmi Múzeumban az a hároméves NKFIH-projekt, amely Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadását tűzte ki célul. A feladat kihívást jelentett a Móricz-műhely számára, hiszen a korábbi, papíralapú kiadási gyakorlatra csak részben támaszkodhattak. A múzeumi informatikai lehetőségek, a filológiai problémák és az alkalmazott szoftverek párbeszédéről szóló esettanulmány a projekt első évének problémafelvetéseiről, megoldásairól szól. Nem törekszik teljes áttekintésre, hiszen munkafolyamat közben ad hírt egy formálódó gyakorlatról.

Kulcsszavak:

digitális filológia, Móricz Zsigmond, kritikai kiadás, levelezés, DigiPhil



1. Bevezetés

A Petőfi Irodalmi Múzeumban (PIM) a tudományos igényű szerzői szövegkiadásoknak hagyománya van, a múzeum többek között Móricz Zsigmond életművének számos forrását, naplóját jelentette meg az elmúlt évtizedben. A modern magyar irodalom más klasszikusainak jelenleg folyamatban lévő kutatásait, posztumusz szövegkiadását tekintve Móricz Zsigmond levelezésének szisztematikus filológiai feltárása is régóta elvégzendő feladat. Ennek érdekében fontos előrelépés történt 2016-ban, amikor megkezdődött a Móricz-levelezés kritikai kiadása az NKFIH támogatásának köszönhetően.¹ A levelezés nagyságrendjét jelzi, hogy a hagyatékban (a PIM különgyűjteményében) található, Móricz-nak címzett levelek, illetve a rokonoknak írt vagy másolatban

¹ Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal (NKFIH), szerződés nyilvántartási száma: 116201.

fennmaradt szerzői levelek száma 9540 darab. A kiadásnak ugyanakkor eleve számolnia kell a levelezéskorpusz önfeltáró jellegével, s a forrásfeltárás a szövegleírással párhuzamosan ma is folyik. A pályázat beadásával egy időben zajlott a PIM-ben a digitális kritikai kiadásnak mint szolgáltatásnak a továbbfejlesztése, így az NKFIH-projekt már beadásakor tartalmazta a DigiPhil digitális szövegkiadási műhely tapasztalatait, integrálta fejlesztési ambícióit. A digitális közzététel a Móricz-kutatásban kísérleti módszernek számít.²

A projekt során olyan új típusú feladatokat kellett megoldani, melyekhez a kutatócsoport hagyományos és digitális filológiai tudását egyesíteni kellett. A feladat a digitális műhely részéről is kihívásnak számított, hiszen a DigiPhil addig csak könyvalapú kritikai kiadások online közzétételének specifikációjával foglalkozott. A szoftverkörnyezet a DigiPhil-ben már alkalmazott komponensekre épül, a feldolgozás módszereit ehhez igazítva alakítottuk ki. Móricz Zsigmond leveleinek digitális kiadása több fázisú munkafolyamatként indult el: fő kérdése az volt, hogyan tudja megteremteni saját munkafolyamatának és a digitális kritikai kiadásnak egymással összekapcsolódó komplex informatikai hátterét. A múzeumi informatikai lehetőségek, a tisztán filológiai problémák és az alkalmazott szoftverek párbeszédéről szóló esettanulmány a projekt első évének problémafelvetéseiről, megoldásairól szól, munkafolyamat közben ad hírt egy formálódó gyakorlatról.

2. A projekt megvalósulásának lépései

2.1. A levelek digitalizálása

A Móricz-levelezéskiadás kéziratkatalógus hiányában a *Huntéka-M* könyvtári-múzeumi integrált rendszerre³ támaszkodik, amely a hasonló projektek esetében még nem gyakori eljárás. A PIM gyűjteménykezelő rendszere⁴ tartalmazza a PIM-ben található kéziratok strukturált alapadatait (levélíró, címzett, dátum, terjedelem, állapotleírás, nyelv). A 2676 szerzői és 6864 írónak címzett levél metaadatainak exportálása után táblázatos listák készültek, így egy felületen vált megjeleníthetővé a levelezés időrendje.

Első lépésben a hároméves projekt során kiadni kívánt, 1923-ig terjedő időszak leveleinek másolatait állítottuk elő e lista alapján: Móricz Zsigmondnak szóló 1582 levelet, Móricz 314 levelét és 788 keltezetlen levelet. A válogatás alapján 2684 kézirat digitalizálását végeztük el. Móricz Zsigmond leveleinek 93 százaléka még kézírással készült, ugyanez elmondható a Móricznak címzett levelekről is. A fakszimilék jelentős segítséget adtak az átiratok elkészítéséhez, megfeleltek az archiválási és állományvédelmi szempontoknak, hiszen az eredeti dokumentumokat nem kellett újra átmozgatni az ellenőrzési munkafolyamat során. 2017-ben magángyűjteményekben is folytattuk a feltárásokat, több száz ismeretlen levélről a helyszínen készítettünk digitális másolatokat. A kutatás során feltárt új források a feldolgozandó korpusz darabszámát jelen-

² DigiPhil: A magyar irodalomtudomány filológiai portálja, hozzáférés: 2018.04.12, <http://digiphi1.hu/>.

³ *Múzeumi Huntéka*, hozzáférés: 2018.04.12, <https://qulto.eu/muzeumi-hunteka>.

⁴ Petőfi Irodalmi Múzeum Huntéka-M online felülete, hozzáférés: 2018.04.12, <https://opac.pim.hu/>.

tősen megnövelték, ezért a kutatásütemezést át kellett dolgozni. A Móricz-levelezés kritikai kiadása az 1913-as évvel zárul, ebből az évből 1 225 dokumentum ismert.

A mesteranyag tárolási formátuma a TIFF (Tag Image File Format), mivel veszteségmentesen tömörít, és alkalmas a képek metaadatainak tárolására is. A mesterfájlok biztonságos elhelyezéséről a PIM informatikai rendszere gondoskodik. A levelek digitális faksimiléi a szolgáltatás keretein belül megtekinthetők lesznek, a digitális kiadás szövege mellett található ikonra kattintva az eredeti forrás digitális másolata összevethetővé válik az átirattal. A kéziratok darabszámához képest sokkal több, közel 7000 képfájl készült: a képeslapok és borítékok mindkét oldaláról, a levelek üres oldalairól is.

2.2. Szövegkódolás

A filológiai szempontokat tekintve legfontosabb feladat a metaadatok leírása és a szövegjellemzők azonosítása, a céloknak megfelelő jelölőnyelv alkalmazása. A Móricz-kutatócsoport a TEI (Text Encoding Initiative) XML (Extensible Markup Language) ajánlása⁵ mellett döntött, alkalmazkodva ezzel a DigiPhil korábbi gyakorlatához, amely megfelel a nemzetközi elvárásoknak. Elsődleges szempont volt a szövegekben előforduló sajátosságok szofisztikált jelölése, a kritikai kiadás filológiai alapvetése szerint. A projekt TEI elemkészletét a DigiPhil alakította ki a kutatócsoport igényeinek megfelelően, a TEI P5 kéziratleírásra kidolgozott modulja szerint (Manuscript Description).⁶

A levelezéskiadások általában csak a levél szövegét és annak hordozóját írják le, nem foglalkoznak például a borítékkal vagy a mellékletekkel, pedig a mellékelt rajzok vagy versek relevánsak a korpusz egészének szempontjából. A boríték, a levél és a melléklet egy objektum egyes részeinek tekinthetők, leírásuk ennek a sorrendnek megfelelően történik. A kutatócsoport nemzetközi digitális levelezéskiadások gyakorlatában nem talált példát ilyen típusú leírásra, ugyanakkor meg tudtuk oldani, hogy a TEI-fájlban ez a három egység a metaadatok és a szövegkódolás szintjén egyaránt reprezentálható legyen.

Moduláris felépítése miatt a TEI kódolásában mindez egy fájlban belül is megoldható a metaadatok szintjén:

```
<msPart style="envelope">
<msPart style="letter">
<msPart style="attachment">
```

Hasonlóan a szövegleírás szintjén is:

```
<div type="envelope" style="handwritten">
<div type="letter" style="handwritten">
<div type="attachment" style="handwritten">
```

⁵ „TEI: P5 Guidelines,” Text Encoding Initiative, hozzáférés: 2018.04.12, <http://www.tei-c.org/Guidelines/P5/>.

⁶ „Manuscript Description,” P5: Guidelines for Electronic Text Encoding and Interchange 3.3.0, 2018. jan. 31., hozzáférés: 2018.04.12, <http://www.tei-c.org/release/doc/tei-p5-doc/en/html1/MS.html>.

A fenti hármas struktúrán túl a kéziratszöveg átírásának linearitása további problémákat vetett fel. Például idegen szöveg beékelődése esetén az időrend volna mérvadó, ám ez sokszor nem rekonstruálható egyértelműen. Ilyen esetben, vagyis, ha a szöveg-szegmentumok sorrendje vitatható, a sorrend meghatározásakor a levél struktúráját követjük, olyan szempontokat is figyelembe véve, mint az olvasás iránya.

A levéltől részben elkülönülő paratextuális szövegszegmentumok, mint a postabélyegző vagy a pecsét, alapvető információkat tartalmazhatnak. A kritikai kiadás objektumstruktúrája így a következőképpen alakult: boríték, levél (ezen belül: fejléc, nyomtatvány, rajz, lábléc, pecsét), melléklet. A kéziratban található szövegszegmentumok elhelyezkedését nem szükséges jelölni a fenti struktúrán kívül, ezekről ugyanis a kéziratokról készült faksimile nyújt információt.

A jelölőnyelvi leírás számos szövegjellemzőt rögzít (pl. betoldás, aláhúzás, javítás, idegenkezűség). Az elírásokat, betűkimaradásokat és -tévesztéseket, az értelemzavaró helyesírást, a nehezen értelmezhető rövidítések feloldását, régi szavak, szóalakok rövidítéseit a könnyebb olvashatóság érdekében a betű szerinti átirat megtartása mellett emendáltuk. Az online kiadásban a betűhű átirat és az olvasószöveg egyaránt olvasható lesz.

A levélküldés folyamata előtti (pl. használt papírra írt) és utáni rájegyzéseket nem a főszövegben, hanem a szerkesztői jegyzetben tesszük közzé: a kézirattest elő- vagy utóélete részének tekintjük, s mint a levél formai jellemzőjét a levélszöveg metaadatai között szereplő levélleírásban szerepeltetjük. Jellegetes példa Pallagi Gyula Móricznak szóló 1900. november 18-a után keletkezett levelén Móricz *A szép lány suttog...* kezdetű, mindeddig ismeretlen *Pua* címmel emlegetett versének két versszaka.⁷

Ahogy más projektek esetében láthattuk, az egyes kiadások a TEI-ajánlásokat alapul véve saját sémákat hoznak létre.⁸ A Móricz-levelezés kódkészleténél is több egyedi megoldás született. Ezek közül egyetlen példát szeretnénk kiemelni, amelynek bevezetése a levélkéziratokon található szövegszegmentumok bonyolult felépítésének köszönhető.

A TEI logikája az egyes szövegszegmentumok elkülönítésére több megoldást tesz lehetővé. A Móricz-kiadásban például az idegenkezűség jelölése eltér a nemzetközi gyakorlattól: idegenkezűség esetén általában a <handShift/>, illetve az <anchor/> jelölőt alkalmazzák. Ezek a jelölők azonban pontszerűek, így ezeknek a szegmentumoknak a kiemelése és vizualizálása informatikai szempontból komoly nehézséget okozott volna. A <seg> jelölő használatával ezt könnyedén elkerülte a kutatócsoport, anélkül, hogy megsértette volna a TEI ajánlását. A @corresp attribútum segítségével lehet megadni a levélre rájegyző nevét, így ezek az információk összekapcsolódnak és kereshetővé válnak, a @type attribútumban pedig a főszövegben található szegmentum jellegét (pl. titkosírás, pecsét) lehet definiálni.

```
<p><seg type="handShift" corresp="Móricz Zsigmond"></seg></p>
```

⁷ Pallagi Gyula levele Móricz Zsigmondnak, Budapest, 1900. november 18. után, PIM Kézirattár, M. 130.

⁸ A Vincent Van Gogh-levelezés kiadás készítésekor például kiegészítették a TEI-sémát saját jelölőkkel. Leo Jansen, Hans Luijten and Nienke Bakker, eds., *Vincent van Gogh – The Letters*. Version: December 2010. Amsterdam & The Hague: Van Gogh Museum & Huygens ING., http://vangoghletters.org/vg/about_6.html.

A <seg> elem nemcsak az idegenkezűség jelölésére szolgál; így jelölendő az aláírás (<seg type="signature">Dr. László</seg>) és a pecsét (<seg type="stamp">).

Szintén a TEI-ajánlástól eltérő megoldás a levél zárlatának jelölése. A TEI a <closer> címkét javasolja, azonban a szintaxis alapján ezt más szövegrész nem követhetné. A problémát a kutatócsoport úgy oldotta meg a <closer> kihagyásával, hogy az aláírást követő szövegrészeket, mint például a lábléc (<floatingText type="footer">), a TEI által szorosan nem definiált szövegszegmentumok leírására szolgáló címkével jelölte.

2.3. A TEI-fejléc lehetőségei

A kéziratra vonatkozó adatokból a projekt szempontjából releváns metaadatok körét a kutatócsoport állapította meg. A kritikai kiadás jellegéből fakadóan a formai leíráshoz használt mezőkészlet jóval gazdagabb, mint az a könyvtári bibliográfiai feldolgozásnál megszokott. Így például három külön adatelem a megírás helye, a feladás helye és az átvétel helye, de ugyanez érvényes a dátumot leíró mezőkre is, a kutatócsoport megadja a megírás, a feladás és az átvétel dátumát. A múzeum könyvtári adatbázisának *kézirat* űrlapján a hely és dátum leírásához csak a *keletkezés helye* és *ideje* HUNMARC-mezők (ismételhető c260\$a és c260\$c) állnak rendelkezésre.

A TEI-ajánlás szerint az XML-fejléc része részletes metaadat-rögzítésre ad lehetőséget, sőt lehetőséget nyújt adatgazdagításra is. Az információk hozzáadása a metaadatokhoz egyrészt saját erőforrásból, a háttéradatbázisok segítségével történik, másrészt külső tezauruszok, névterek bekapcsolásával. Az adatgazdagítás négy fő dimenziója (személy, hely, idő, fogalom) közül jelenleg a személy, a hely és a fogalom vonatkozásában történik bővítés.

A TEI-fejléc nyitó és záró címkéje közötti rész tartalmazza az objektum metaadatait. A <teiHeader> alatt a <fileDesc> foglalja össze a digitális kiadásra vonatkozó információkat. A <title> elemen belül a kiadás címét:

```
<titleStmt>
  <title>Móricz Zsigmond levelezés kritikai kiadás</title>
</titleStmt>
```

A <publicationStmt> rész jól reprezentálja a finomítási lehetőségeket. A neveket jelölő elemeken belül megadható, hogy személyről, intézményről vagy helységnévről van szó (<persName>, <orgName>, <placeName>).

Itt már látható példa az adatgazdagításra is URI-k megadásával a <ref> elem @type attribútumában: a kiadó nevéhez bekerült a PIM VIAF⁹ katalógusában lévő azonosítója, valamint a kiadás helyénél a GeoNames¹⁰ egyedi azonosítója. Hasonlóképpen hivatkozik az <availability> címke a közzétételi jogokra.

Nagyon fontos megadni a feldolgozott objektum perzisztens egyedi azonosítóját (PID) és URI-ját (Uniform Resource Identifier), ugyanis ezek az azonosítók garantálják az egyes digitális objektumok (jelen esetben a levelek) azonosíthatóságát és a kiadás idézhetőségét. Ezeket szintén a @type attribútum jelöli az <idno> elembe.

⁹ Virtual International Authority File, hozzáférés: 2018.04.12, <http://viaf.org/>.

¹⁰ GeoNames, hozzáférés: 2018.04.12, <http://www.geonames.org/>.

```
<publicationStmt>
  <publisher>
    <orgName>Petőfi Irodalmi Múzeum</orgName>
    <ref type="url">http://viaf.org/viaf/152132060</ref>
    <ref type="url">http://www.pim.hu</ref>
  </publisher>
  <pubPlace>Budapest <ref type="url">http://www.geonames.org/
    3054643</ref>
  </pubPlace>
  <date>2015</date>
  <availability>
    <p>©Free Access - no-reuse <ref type="url">http://www.europeana.eu/
rights/rr-f/</ref>
    </p>
  </availability>
  <idno type="PID">o:PKEL.M.100-2553-18_a</idno>
  <idno type="URL">o:PKEL.M.100-2553-18_a</idno>
</publicationStmt>
```

A dokumentum egészének leírása a <sourceDesc> elem alatt található további strukturált egységekben. Az <msDesc> (manuscript description) címke jelöli a kézirat metaadatainak leírására vonatkozó információkat. A lelőhely megadásánál az <msIdentifier> elemnél lehetne leírni a provenienciára vonatkozó információkat, de mivel a közgyűjteményekben ezek érzékeny adatoknak számítanak, a TEI-kódban és a DigiPhil oldalán az adatok nem lesznek nyilvánosak, egyelőre csak a hagyaték neve szerepel az <msName> alatt felvéve.

```
<sourceDesc>
  <msDesc>
    <msIdentifier>
      <country>Magyarország</country>
      <settlement>Budapest<idno type="KOHA_GEO">KOHA_GEO:9227</idno>
      </settlement>
      <institution>Petőfi Irodalmi Múzeum</institution>
      <repository>Petőfi Irodalmi Múzeum Kézirattár</repository>
      <idno>PIM M. 100/2553/18</idno>
      <msName> Móricz Zsigmond-hagyaték </msName>
    </msIdentifier>
```

Az objektum egyes elemeit az <msPart> elem szegmentálja. A levél fizikai leírását a <physDesc> címke vezeti be, melynek további részeivel megadható a levél mérete és állapota.

```
<msPart style="letter">
  <msIdentifier/>
  <physDesc>
```



```

<objectDesc>
  <supportDesc>
    <extent>
      <measure type="quantity" unit="folio"> 4 </measure>
      <dimensions unit="mm">
        <height> 109 </height>
        <width> 174 </width>
      </dimensions>
    </extent>
    <condition>
      <p>Sárgult papíron.</p>
    </condition>
  </supportDesc>
</objectDesc>
</physDesc>
</msPart>

```

A levél részletes leírása a <profileDesc> címke alatt látható. Ebbe a részbe került adatgazdagítás céljából egy formai tárgyszó a Getty Art & Architecture (AAT)¹¹ teauruszából. A TEI-fejlécnek ebben a részében található a megírásra, a feladásra és az átvételre vonatkozó metaadatok feltüntetése a <creation>, <correspAction type="sent"> és a <correspAction type="recieved"> jelölők segítségével.

```

<profileDesc>
  <langUsage>
    <language ident="hu"/>
  </langUsage>
  <textClass>
    <keywords scheme="AAT" corresp="Letter">
      <term>levél</term>
      <idno type="AAT"> AAT:300026879 </idno>
    </keywords>
  </textClass>
  <creation>
    <date when="1905-07-02"/>
    <placeName>Budapest <idno type="KOHA_GEO">KOHA_GEO:9227</idno>
  </placeName>
  </creation>
  <correspDesc>
    <correspAction type="sent">
      <persName>Móricz Zsigmond <idno type="KOHA_AUTH">KOHA_AUTH:120256
    </idno>
    </persName>

```

¹¹ „Art & Architecture Thesaurus,” The Getty Research Institute, hozzáférés: 2018.04.12, <http://www.getty.edu/research/tools/vocabularies/aat/>.

```

</correspAction>
<correspAction type="recieved">
  <persName>Holics Janka <idno type="PIM">PIM:1153120</idno>
</persName>
</correspAction>
</correspDesc>
</profileDesc>

```

2.4. Névterek

Az egységes besorolási adatok (személy- és helynevek) a szolgáltatás hozzáférési pontjait biztosítják és a szemantikai kapcsolatok kiépítését segítik. A Móricz-levelek átírásakor jelenleg a személyek, a földrajzi helyek és a műcímek azonosítását végzik a munkatársak. Az identifikáció része, hogy az entitások egyedi, állandó azonosítót kapnak, így az összes előfordulásuk kereshetővé válik.

```

090 __ a INT
a IPA
a ITO
a IKN
a TLA
a GYN
100 1_ a Móricz
j Zsigmond
d 1879-1942
400 __ Q adateltérés[$j]:Zsigmond (?)
400 __ a Zsiga
j bácsi
400 __ a M.
j Zs.
500 1_ a Holics
d 1883-1925
j Janka
667 __ a író
667 __ a újságíró
667 __ a publicista
667 __ a lapszerkesztő
680 __ a sírhely: 34. parcella
900 __ a 1879. VI. 29.
902 __ 3 ITE-9395
a Csécse

```

1. ábra. Móricz Zsigmond besorolási rekordjának HUNMARC mezői a Huntékában
Forrás: *Huntéka-M*, PIM

A PIM személynévtér-állománya hozzávetőlegesen 600000 rekord.¹² Az életrajzi típusú adatbázisok közül a *Magyar Életrajzi Index*¹³ rendelkezik érvényes, kontrollált személynév-rekordokkal, ezért a projekt számára a Huntéka-rendszerből ezt a részt migrálták a DigiPhil mögött működő könyvtári rendszerbe, a *Kohába*.¹⁴ A PIM névtérben lévő névrekordok sok hozzáadott információt tartalmaznak (rokoni kapcsolatok, lakhelyek, temetés helye stb.), így a megfeleltetésnél redukálni kellett a kiegészítő adatok körét. Csak annyi mező került át a *Koha*-névrekordokba, amennyi minimálisan elegendő ahhoz, hogy egy személy azonosítható legyen: vezetéknév, keresztnév, a születés és halálozás adatai.

Authority #120256 (Personal Name)

Used in 111 record(s)

0	1	4	5	6
000 - LEADER @ 00494n a2200229 i 4500 001 - CONTROL NUMBER @ 120256 003 - CONTROL NUMBER IDENTIFIER @ 65715 005 - DATE AND TIME OF LATEST TRANSACTION @ 20170412090807.0 008 - FIXED-LENGTH DATA ELEMENTS @ 080712s 1 040 ## - CATALOGING SOURCE a Original cataloging PIM c Transcribing agency DigiPhil 902 ## - Születési hely a Születési hely Csécse 906 ## - Halálozási hely a Halálozási hely Budapest				
100 1# - HEADING--PERSONAL NAME a Personal name Móricz Zsigmond d Dates associated with (1879-1942)				

2. ábra. Móricz Zsigmond besorolási rekordja a *Kohában*. Forrás: <http://biblio-intra.digiphil.hu/>

A Móricz-projekt a PIM és a *Koha* azonosítóit használja a személynevek egyértelműsítésére.

```
<persName>Édesapám<idno type="PIM" corresp="Móricz Bálint">PIM:
297674</idno></persName>
```

Vannak olyan esetek azonban, amikor nincs elegendő adat a személy azonosításához, és így nem lehet érvényes névrekordot létrehozni az adatbázisban. A levelek irodalmi, művészeti, közéleti kapcsolatokat felvonultató adatai mellett családi vagy személyes levélváltásokra is nagy mennyiségben van példa a Móricz-levelezésben. A rokonok, barátok leveleiben sokszor előfordul csak keresztnévvel említett személy, például a

¹² Bánki Zsolt, Mészáros Tibor, Németh Márton és Simon András, „Azonos személyekre vonatkozó név besorolási rekordok automatikus felderítése a PIM adatbázisában,” *Tudományos és Műszaki Tájékoztatás* 63, 12. sz. (2016): 471.

¹³ *Magyar Életrajzi Index*, hozzáférés: 2018.04.12, <https://opac-nevter.pim.hu/>.

¹⁴ *Koha Library Software*, hozzáférés: 2018.04.12, <https://koha-community.org/>.

Móricz-háztartásban segédkező alkalmazottak (Anna cseléd) vagy Móricz testvérének osztálytársai (pl. Sanyi). Ezek az entitások nem kerülnek be az adatbázis besorolási állományába, de a kritikai kiadásban fontos a megkülönböztetésük, ezért ún. lokális azonosítót (LOK) kapnak.

```
<persName>Anna<idno type="LOK" corresp="Anna_cseléd">LOK:00013
</idno></persName>
```

A helynevek és a műcímek identifikációja hasonló módon történik. A helynevek azonosításához a Geotaurusz¹⁵ rekordjait importáltuk a *Kohába*, így a levelek szövegeiben lévő helynevek egyedi azonosítói a *Kohából* kerülnek a TEI-be.

```
<placeName>Gödöllőig<idno type="KOHA_GEO"
corresp="Gödöllő">KOHA_GEO:21799</idno></placeName>
```

3. Tervek

3.1. Kommunikáció integrált rendszerekkel *Huntéka-M*, *Koha*

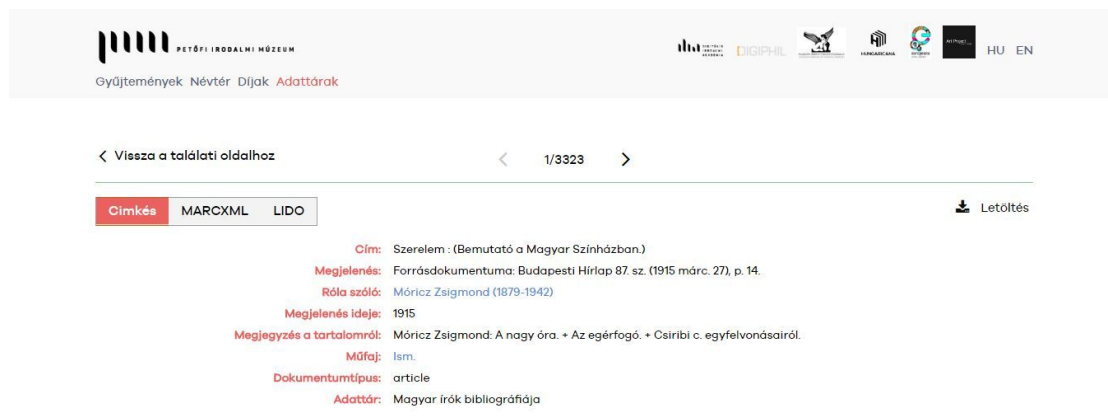
A kéziratok feldolgozásának alapja a levelek formai feltárásának elvégzése, és a szövegekben előforduló entitások (a személyek, a földrajzi helyek és a címek) azonosítása.

A *Huntéka-M* rendszerében nemcsak a gyűjtemények anyaga található, hanem a PIM tevékenységéhez tartozó háttérkutatások eredményei is. A múzeum jelentős személynévtérrel rendelkezik, de emellett egyéb bibliográfiai és faktográfiai jellegű adatbázisokat is épít. A különböző forrásokból érkező heterogén adatok egy integrált rendszerbe migrálásával a *Huntéka-M* már nemcsak a múzeumi nyilvántartás funkcióit látja el, hanem szakirodalmi tudásbázisként képes kiszolgálni a kutatói igényeket is. A projekt szempontjából különösen jól használható a *Magyar írók bibliográfiája*¹⁶ és a *Budapest topográfia*¹⁷ az első számos Móricz-vonatkozású cikk leírását és forrásait tartalmazza, a másodikban pedig nyomon követhetjük Móricz Zsigmond budapesti lakcímeit.

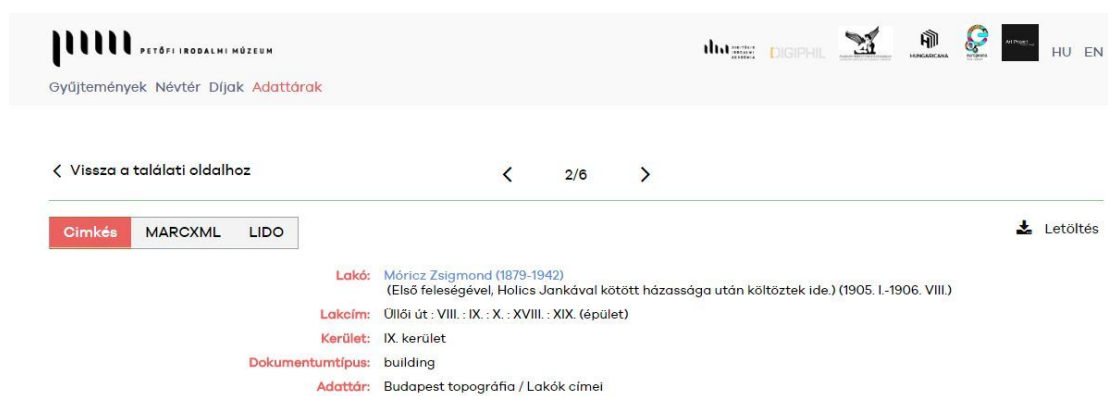
¹⁵ Ungváry Rudolf és Cserbák András, szerk., „Geotaurusz és Geohistaurusz: Földrajzi nevek és humángéográfiai nevek tezaurusza,” 2001. nov. 1., hozzáférés: 2018.04.12, <http://mek.oszk.hu/00000/00070/html/>.

¹⁶ *Magyar írók bibliográfiája*, hozzáférés: 2018.04.12, <https://opac-adattar.pim.hu/>.

¹⁷ *Budapest topográfia*, hozzáférés: 2018.04.12, <https://opac-nevter.pim.hu/>.



3. ábra. Egy Mórícz-vonatkozású cikk rekordja a *Magyar írók bibliográfiája* adatbázisban. Forrás: <https://opac-adattar.pim.hu/record/-/record/PIM1367718>



4. ábra. Mórícz Zsigmond lakcímének rekordja a *Budapest topográfia* adatbázisból. Forrás: <https://opac-adattar.pim.hu/record/-/record/PIM1644715>

A *Huntéka-M* szabványos kimenettel és szabványos adatcsere-formátummal (HUN-MARC) rendelkezik, ezért könnyen kommunikál más integrált rendszerekkel, így a DigiPhil bibliográfiai és besorolási adatait tároló *Kohával* is. A biblio.digiphil.hu mögött működő könyvtári komponens nyílt forráskódú, amint a DigiPhil más célszoftverei is. A *Kohában* épülnek a Mórícz-levelezéshez tartozó elsődleges és másodlagos bibliográfiák, és ide integrálódnak a más rendszerekből érkező besorolási állományok.

A Mórícz-levelezés a bibliográfiai rekordok közül a kézirat, könyv, periodika, cikk, a besorolási rekordok közül pedig a személynév, földrajzi név, egységesített cím űrlapjait használja majd.

Az adatbázisban viszont problémát jelent egy elvi mű és a kiadások kapcsolatainak leképzése a MARC korlátozottsága miatt. A kritikai kiadás mellett épülő bibliográfiában egy adott mű rekordjában jelenne meg az is, ha egy regényből átdolgozás (pl. színdarab) készült, ahogy a Mórícz által írt művek esetében ez többször előfordult. Szintén nehéz MARC-sémával leírni, amikor a levélben csak általánosan említenek egy művet, és nem egy konkrét kiadásról van szó, vagy amikor a mű címe csak

ötletként merül fel, de később nem íródott meg. MARC-ban az egységesített cím (a130) besorolási rekord almezőibe nem lehet elhelyezni a szerzőséget, azt csak a kapcsolódó, a mű kiadásait leíró bibliográfiai rekordok mutatják (c100 – Személynév főtétel). A megoldást a könyvtári világban egyre inkább teret hódító FRBR-alapú RDA (Resource Description and Access) katalogizálási szabályzat¹⁸ jelentheti. A bibliográfiai tételek funkcionális követelményeit (FRBR) megfogalmazó entitáskapcsolat-modell külön értelmezi egy mű kifejezési formáját (*expression*), megjelenési formáját (*manifestation*) és példányát (*item*).¹⁹

A Koha wiki oldalán²⁰ láthatjuk, hogy a közösség fejlesztői már kidolgozták, hogyan igazítható a MARC-alapú rendszer az RDA igényeihez, és vannak már olyan projektek, amelyek sikeresen implementálták az FRBR rendszerét a Kohába. Valószínűleg a DigiPhil előtt álló egyik fejlesztési feladat a projektet kiszolgáló adatbázis felkészítése lesz az RDA-alkalmazásra.

A Koha előnye, hogy rendelkezik Z39.50 protokollal, így képes más adatbázisokból rekordokat fogadni. A kritikai kiadáshoz a HUMANUS²¹ Móricz-vonatkozású cikkei kerültek az adatbázisba, ahol az átvett rekordok 040-es mezője mutatja az eredeti forrást.

A Móricz-projekten belül a Kohában tárolt bibliográfiai rekordoknak kettős funkciója lesz: egyrészt segítik a kutatási munkát, másrészt a DigiPhil oldalán tájékoztatnak a művek és a kéziratok metaadatairól. A kutatás feladatának tekinti a Móricz-bibliográfia építését – többek között ezért is értelmezik egységesített címként a szövegekben előforduló Móricz-műcímeteket. Ha konkrét kiadásról vagy példányról esik szó egy levélben, akkor a szövegkódolásban ezt szintén jelzik.

3.2. Keresés

A metaadatokba és a szövegekbe illesztett egyedi azonosítók (vagyis az összetett struktúrájú TEI-elemkészlet alkalmazása) többfunkciós kereséseket tesznek majd lehetővé az infrastruktúráját kiaknázva a digitális kiadású DigiPhil Móricz-levelezésben.

A DigiPhil a kutatás során új keresőfelületet fejleszt, amely ötvözi a szabad szavas keresést és az XML-nyelv adta lehetőségeket. Egy indexelő alkalmazás a korpusz szövegének egészét feldolgozza, a szabad szavas keresésen túl lehetőség nyílik a csonkolt szavak és az ún. *joker* karakterek alkalmazására is. A keresőfelület másik oldalán az XML-ek hierarchiáját és elemkészletét kezelő eszköz áll. Ennek segítségével az egyes TEI-elemekre külön-külön is lehet keresni (például a törölt szövegrészekre:); illetve különböző szűrési feltételeket lehet majd beállítani, így például ha dátumra keres a felhasználó, előre megadható lesz, hogy a háromféle datálást milyen sorrendben vegye figyelembe a keresőrendszer.

¹⁸ „Resource Description and Access (RDA),” Library of Congress, hozzáférés: 2018.04.12, <http://www.loc.gov/aba/rda/>.

¹⁹ „Functional Requirements for Bibliographic Records,” IFLA, hozzáférés: 2018.04.12, https://archive.ifla.org/VII/s13/frbr/frbr_current_toc.htm.

²⁰ „Koha: RDA,” hozzáférés: 2018.04.12, <https://wiki.koha-community.org/wiki/RDA>.

²¹ Humántudományi Tanulmányok és Cikkek Adatbázisa, hozzáférés: 2018.04.12, <http://www.oszk.hu/humanus/>.

4. A projekt összegzése

Az elmúlt másfél évben a kutatócsoport a projekt sikeres megvalósításához a szükséges és nélkülözhetetlen alapokat rakta le. Megtörtént a lelőhelyek feltérképezése, a kiadás szempontjából releváns, jelenleg elérhető forrásanyag számbavétele. Rendelkezésre állnak a PIM gyűjteményében található Móricz-levelezés példányairól készült fakszimilék, a képek szabályos elnevezése, ezen túlmenően a képszerkesztés folyamata befejeződött. A más közgyűjteményekben található levelek digitális másolatainak megrendelése folyamatban van.

Elkészült az 1913-ig keletkezett kéziratok főszövegeinek leírása (közel 1300 levél): a szöveggkritikai elvek szerinti betűhív átiratok, valamint az emendálásokat tartalmazó olvasószöveg előállítás is. A szövegek ellenőrzése, összeolvasása, szoros időrendbe rendezése folyamatosan halad. A szerkesztőbizottság kialakította a levelekben előforduló szövegjellemzők és szövegműveletek jelölésére használt TEI-elemkészletet. A definiálni kívánt entitások azonosítása a személyneveket, földrajzi neveket és az egységesített címeket érinti, amelyek mindegyikéhez egyedi azonosítókat rendelnek a kereshetőség és a szemantikus kapcsolatok kiépítésének érdekében.

A munkatársak elvégezték a levelek formai feltárását, vagyis a kéziratok fizikai adatainak felvételét, valamint a szöveggközlést a kritikai kiadásnak megfelelő részletezéssel.²² A kiadáshoz használt besorolási adatok tárolása és újabb rekordokkal való bővítése a *Koha*-rendszerben történik, ahol még a rekordstruktúra és a rekordkapcsolatok rendszere folyamatos fejlesztés alatt áll.

A projekt felénél, 2017-ben már számos tanulságot vonhatott le a kutatócsoport a munkafolyamatok eredményességét illetően. A szöveggfeldolgozás egyes fázisainak munkamódszerei változó hatékonyságúnak bizonyultak. Bár a levelek szkennelése, a szövegek metaadatolása és *Microsoft Word*ben való jelölése megfelelően haladt, mára bebizonyosodott, hogy a köztes platform használata túl sok hibalehetőséget rejt a szövegtranszformáció során. A *Microsoft Word*-fájlokból nem lehet egy lépésben TEI XML-fájlokat kinyerni, csak bonyolult, többlépcsős folyamattal. A *Microsoft Word*ben történő átalakítás Visual Basic-kóddal (Visual Basic for Applications)²³ és reguláris kifejezések segítségével történik, majd a kinyert (még nem hierarchikus) XML-fájlokat *Oxygen XML Editor*ben²⁴ alakítják TEI XML-kóddá a projekt számára írt egyedi stíluslap segítségével. A *Microsoft Word*ből való átmásolás nehézségekkel terhelt a karakterkódolás miatt is (idézőjelek, rövid és hosszú kötőjelek keveredése), nem beszélve a szövegbevitel és a kódolás során történő hibás jelölésekről, gépelési hibákról. A többféle ellenőrzőprogram futtatása, a hibajavítások, az újabb ellenőrzések beiktatása mind jelentősen megnöveli a szövegtranszformációra fordított időt, és fennáll az adatvesztés veszélye. A projekt következő szakaszában a szövegek kódolása csak szabványos TEI XML-környezetben történhet a célnak megfelelő eszközzel. A

²² Magyar Tudományos Akadémia I. osztályának Textológiai Munkabizottsága, „Alapelvek az irodalmi szövegek tudományos kiadásához,” hozzáférés: 2018.04.12, <http://textologia.iti.mta.hu/alapelvek.pdf>.

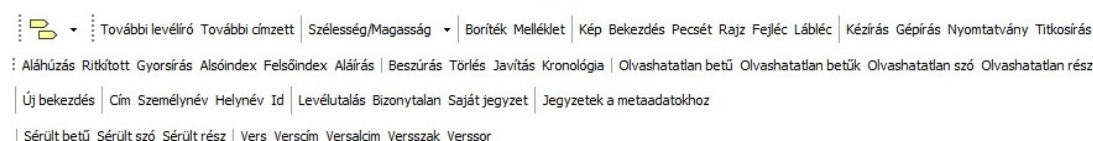
²³ „Visual Basic Guide,” Microsoft, hozzáférés: 2018.04.12, <https://docs.microsoft.com/hu-hu/developer/visual-basic/>.

²⁴ *Oxygen XML Editor*, hozzáférés: 2018.04.12, <https://www.oxygenxml.com/>.

hibák kiküszöbölésére a DigiPhil csapata új leírókörnyezetet fejlesztett az *Oxygen XML Editor* programban, amely felváltotta a *Microsoft Word*-öt mint adatbeviteli felületet.

A DigiPhil a virtuális kutatókörnyezetet²⁵ először az Arany János levelezése kritikai kiadásainak digitalizálási projektjén tesztelte, amely az Arany János Összes Művei 15–19. kötetekben található. A sikeres próbaidőszak után a Móricz-kutatócsoport is áttért az új környezet használatára a levelek leírásához. A kutatókörnyezet a *Microsoft Word*-del szemben számos előnnyel rendelkezik. Azon túl, hogy szabványos kimenetet biztosít, számos hibalehetőséget is megelőz. A kutatókörnyezet szintaktikai ellenőrző algoritmusokat tartalmaz, amelyek figyelmeztetnek a formalizálható szintaktikai hibákra. A DigiPhil a *Microsoft Word*-fájlok átalakításából átörököltette a köztes, kevés hierarchiát tartalmazó XML-struktúrát a levelek leírásához, mivel a TEI szerkezete rendkívül bonyolult, ez a struktúra jelentősen megnehezítette volna a kutatói környezet kialakítását, az XML-ek megjelenítését és a szintaktikai ellenőrzést. Ezekből a „sík” (keves hierarchiát tartalmazó) XML-fájlokból a DigiPhil stíluslap segítségével állítja elő a publikálásra szánt TEI XML-fájlokat. A kutatókörnyezet használatához elég az *Oxygen*-szerkesztőt egyszer telepíteni és importálni a leíráshoz fejlesztett komponenseket. Természetesen a levelek leírása során új jelenségek bukkanhatnak fel (például ritkított betűkkel írt szó vagy sérült papír miatt olvashatatlan szavak), amelyek kódolására új jelölőket kell bevezetni, illetve ezeknek a környezetbe való beillesztését a szükséges módosításokkal elvégezni. A kutatói környezet egy másik előnyös tulajdonsága, hogy minimalizálja az adatvesztés lehetőségét, és biztosítja a kutatócsoport számára, hogy különböző munkaállomásokon dolgozzanak, a környezet ugyanis összeköttetésben áll egy változáskövető szerverrel, amelyre csak szintaktikailag helyes fájlokat menthetnek. A DigiPhil meghatározott időközönként archiválja a fájlokat egy repozitóriumban, ahonnan a változáskövető szerver esetleges leállása esetén is visszaállíthatók a fájlok.

A kutatókörnyezet kialakításakor a DigiPhil figyelembe vette azt az igényt, hogy a környezetet felhasználóbarát, irodai szoftvereket imitáló grafikus megjelenítéssel lássa el, amely hasonlít a már megszokott *Microsoft Word*-környezethez (menürendszer, magyar feliratú gombok), a gombok segítségével a megfelelő XML-jelölők automatikusan a kijelölt szöveghelyre kerülnek, így elkerülve a *Microsoft Word*-re jellemző szintaktikai hibákat.



5. ábra. Magyar nyelvű menüsor a Móricz-kutatócsoport által használt leíró környezetben

A kutatócsoport a metaadatokat egy előre meghatározott mezőkkel rendelkező táblázatban adhatja meg:

²⁵ Palkó Gábor, „A digitális bölcsészet kultúrtechnikái. Virtuális kutatókörnyezetek,” előadás *A humán tudományok és a gépi intelligencia* c. konferencián, Budapest, 2017. november 20.

Azonosító:	PKEL.M.130-pallagiyulato-0011_a		
Leíróhely:	PIM.M.130/pallagiyulato/0011	Proveniencia:	Kiss Ferenc tulajdonából (2005)
Hagyaték:	Móricz Zsigmond-hagyaték		
Levélíró:	Pallagi Gyula		
Id:	Tipus: KOHA_AUTH	121464	
Levélíró testület:			
Megírás helye:			
Id:	Tipus: KOHA_GEO		
Megírás dátuma:	Mikor: 1900-11-18	-tól	-ig: 1900-11-18
Feladás helye:			
Id:	Tipus: KOHA_GEO		
Feladás dátuma:	Mikor:	-tól	-ig:
Címzett:	Móricz Zsigmond		
Id:	Tipus: KOHA_AUTH	120256	
Címzett testület:			
Átvétel helye:			
Id:	Tipus: KOHA_GEO		
Átvétel dátuma:	Mikor:	-tól	-ig:
Nyelv:	hu		

Levél adatai	
Fólió száma:	4
Darabszám:	
Típus:	levél
Leírás:	Fekete tintával a fóliók mindkét oldalán, alján oldalszámozással.
Szélesség:	109
Magasság:	174
Írástípus:	K
Boríték adatai	
Leírás:	
Szélesség:	
Magasság:	
Írástípus:	
Melléklet adatai	
Fólió száma:	
Darabszám:	
Leírás:	
Szélesség:	
Magasság:	
Írástípus:	
Típus:	
Publikáció:	

6. ábra. Móricz-kutatócsoport által kitöltendő táblázat

Bár a kutatócsoport már a *Microsoft Word*ben való leírás során is végzett adatgazdagítást, az automatikus átalakításhoz kidolgozott szintaxis rendkívül bonyolultnak bizonyult, a legtöbb szintaktikai hiba, elgépelés itt adódott, ami jelentős mennyiségű utólagos ellenőrzést és javítást igényelt a szövegtranszformáció elvégzése után.

Személy- és helynevek azonosítása, valamint bibliográfiai adatok kódolása *Microsoft Word*ben:

```
[személy] [@ Móricz Dezső @ KOHA_AUTH:313737] Dezső [személy vége]
[hely@ Gödöllő @] KOHA_GEO:21799 Gödöllőnél [hely vége]
[cím@ Hét krajcár@] KOHA_TITLE:3081083 KOHA_BIBL:40125836 Hét
krajcár[cím vége]
```

A megfelelő elem beillesztése után a kutatókörnyezet automatikusan létrehozza a kitöltendő mezőket, illetve előre kitölti a névterek DigiPhilben használt azonosítóját. A rendszer által javasolt azonosító (KOHA_AUTH) a személynevek esetén legördülő ablakban jelenik meg, a kutatócsoport itt választhat más névteret (PIM, LOK).

7. ábra. Személy- és helynevek azonosítása, bibliográfiai adatok kódolása kutatói környezetben

Az új kutatókörnyezet kialakításán túl módosult a levelek (és a hozzájuk tartozó XML-fájlok) ellenőrzésének folyamata is. Míg a korábbi tervek szerint a *Microsoft Word*-ben kódolt levelek transzformációja után a javítás a TEI XML-fájlokban zajlott volna, a jövőben a DigiPhil a kutatócsoport számára egy olyan tesztoldalt biztosít, amely a leíró környezetben leírt „sík” XML-eket jeleníti meg, az ott megszokott vizualizációval. Ennek a felületnek a segítségével a már leírt levelek eljuttathatók azoknak a szakértőknek is, akik nem vettek részt a levelek átírásában: ők online, a kutatói környezet telepítése nélkül kapcsolódhatnak be a kutatásba. A tesztfelületen történő ellenőrzés után a kutatócsoport a javításokat még a kutatói környezetben végzi el, majd ezt követően alakítja át a DigiPhil a leveleket, és publikálja a hivatalos oldalán.

Arról, hogy milyen lehetőséget nyújt a digitális médium a kritikai kiadás számára, milyen vizsgálati módszereket ajánlhat fel a levélszövegek vizsgálatára, csak a 2019-ben lezáruló kutatási időszak után nyújtható részletesebb, elméleti kérdéseket is érintő összefoglaló.

The Digital Critical Edition of the Correspondence of Zsigmond Móricz (1892–1913): a Case Study

The NKFIH-project that seeks to publish the digital edition of Zsigmond Móricz's (1892–1913) correspondence was launched in 2016 at the Petőfi Literary Museum in Budapest. The project itself has been a huge challenge for the Móricz-research group because they can only partially rely on the earlier paper-based edition. Drawing on the experiences and resolved problems of the first years of the project, this paper focuses on the harmonization and relationship of a museum's programming/IT possibilities, the philological problems and the applied software capacities. As the paper discusses the challenges of an ongoing project, the study does not offer a holistic and comprehensive overview of the entire project, rather a list of problems as encountered along the way and their solutions.

Keywords:

digital philology, Zsigmond Móricz, digital scholarly edition, correspondence, DigiPhil

Biszak Sándor*Arcanum Adatbázis Kft.*

biszak@upcmail.hu

Kokas Károly*SZTE Klebelsberg Könyvtár*

kokas@ek.szte.hu

Budapest Időgép

A történeti térképészet online és digitális újdonsága, új kutatási eszköze a hálózaton a nemrég elindult projekt, a Budapest Időgép (<https://hungaricana.hu/hu/budapest-idogep/>). A szolgáltatás egymásra vetített különböző korú térképekkel és hozzájuk kapcsolt, georeferált levéltári adatok tömegével és fényképekkel együtt mutatja be Budapest újabb kori történelmének geográfiai dimenzióit. A műhelytanulmány néhány példával bemutatja, milyen különleges felhasználásai lehetségesek a *Budapest Időgép*-szolgáltatásnak.

Kulcsszavak:

történeti földrajz, digitális térképek, georeferálás, időgép, Budapest Időgép

A földrajz a történelem kulcsa.¹Sir Halford John Mackinder
(a *Pivot-elmélet* atyja)

A történelem tanulmányozásának legjobban ismert aspektusa az időbeli vizsgálat. Azt is tudjuk azonban, hogy hasonló, – ha talán kevésbé hétköznapi is – alapfeltétele a megértésnek a történelmi terep, a földrajzi tájegységek, térségek historikus ismerete. Régről ismert, hogy a földrajzi adottságok befolyásolják a történelmi eseményeket, akár egész országok vagy birodalmak történelmét is meghatározhatják. A hatás nagyon összetett, így a fejlődést befolyásolja a természeti erőforrások megléte vagy hiánya, de hatalmas hatása van a domborzat, vízrajz, éghajlat stb. milyenségének is. Azonban a lokalitás a történelem apró részleteire is hatással van. Egyszerűen nélkülözhetetlen az, hogy a *mikor* mellett megmondjuk, hogy pontosan *hol* történt az adott esemény. A térképek digitális feldolgozása, az adatokkal összekapcsolt online szolgáltatások megjelenése soha nem remélt lehetőséghez juttat bennünket.

¹ Halford John Mackinder, „The Geographical Pivot of History,” *Geographical Journal* 23 (1904): 421–437, <https://doi.org/10.2307/1775498>.

A történeti kutatások egyre fontosabb (és divatosabb) eleme a mikrotörténelem kutatása, a neves és névtelen életek aprólékos feltárása, dokumentálása és nyomon követése. E kutatások legfontosabb forrásai a névtárak, lakcímtárak (pl. *Magyarország tiszti cím- és névtára* és a *Budapesti Lakcímjegyzék* rendszeresen megjelenő kötetei),² valamint a lakástulajdonosok adatait, nevét tartalmazó telekkönyvi betétek és egyéb források. Emellett sok egyéb forrás lehetséges, melyek közös jellemzője, hogy lakcímmel, helyrajzi számhoz kapcsolódnak.

A *Budapest Időgép*³ víziója az, hogy gyűjtsük össze a legfontosabb, legrészletesebb térképeket, ha lehetséges 20–40 évenként, és ezeken a térképeken helyezzük el a fellelhető adatokat. A projekt kiindulópontjaként az 1975 és 2000 között megjelent *Topográfiai Mutatók* című kiadvány kötetei szolgáltak.⁴ Ebben kötetenként az egyes városrészek (Terézváros, Ferencváros stb.) telekszámait, helyrajzi számait voltak összegyűjtve és rendezve az aktuális utcanévvel, házszámokkal együtt.⁵ Ez a hatalmas adatár volt a *Budapest Időgép* egyik kiindulópontja.

A régi budapesti térképek (és más jellegű térképek) digitalizálása, georeferálása nem új keletű munka hazánkban sem, mindez az Arcanum által évek óta folyamatosan végzett tevékenység. Jól mutatja ezt a MAPIRE-szolgáltatás,⁶ amely mintegy ötven részletes történelmi térkép digitalizált változatát teszi kezelhetővé, igen komplex módon. Az talán érzékelteti a szolgáltatás különlegességét, hogy a teljes Monarchiát lefedő katonai felmérések, a mai Magyarországot, Ausztriát, Horvátországot, Kárpát-alját lefedő kataszteri térképeket teszi hozzáférhetővé. Valójában több százezer nagy felbontású szelvényről van szó. Budapestről egyébként mintegy 30 térképmű található meg, közöttük minden 1945 előtről ismert és a várost telekszíntelen lefedő térképmű.

Georeferálás

Talán érdemes néhány szót ejteni a történelmi térképek georeferálásáról. Ismeretes, hogy a mai értelemben vett modern kori térképek készítése a 18. század végén kezdődött csak el. Az ezen időszaktól kezdve megjelenő térképeken számos tartalmi újítás jelent meg. A földrajztudomány, a műszaki eszközök, illetve a matematika fejlődése megadta az alapot arra, hogy a valóságot valóban leképező vetületeket lehessen szerkeszteni. A domborzati ábrázolásban megkülönböztették a szintvonalakokat is, és megkezdődtek a nagy területet felölelő, részletes topográfiai felmérések.

² *Magyarország tiszti cím- és névtára, 1873–1944*, Arcanum, hozzáférés: 2018.05.22, <https://adtpl.us.arcanum.hu/hu/collection/TisztiCimtar/>; *Budapesti Cím- és Lakásjegyzék, 1880–1928*, Arcanum, hozzáférés: 2018.05.22, https://library.hungaricana.hu/hu/collection/fszek_budapesti_cim_es_lakasjegyzek/


³ A szolgáltatás a Hungaricana-rendszer keretében került publikálásra: <https://hungaricana.hu/hu/budapest-idogep/>.

⁴ Budapest Főváros Levéltára, *Topográfiai Mutatók* (Budapest: Budapest Főváros Levéltára, 1975–2000).

⁵ Ennek lényege, hogy a telkek azonosító adatait az adatbázis készítésekor rendelkezésre álló helyrajzi számoktól visszafejti a legkorábbi – a török alóli felszabadulás utáni első – megismerhető telekadatokig.

⁶ MAPIRE: *Történelmi Térképek Online*, Arcanum, hozzáférés: 2018.05.25, <http://www.mapire.eu/www.mapire.eu>.

Kösterület: Budapest, Terézváros VI. Nagymező utca
Mai cím: Nagymező utca 16.
Helyrajzi szám (1982): 29337
1879, helyrajzi szám: 3916
1879, cím: Nagymező utca 16.
1876 körül, Telekösszeírás, VII. füzet: 1187
1862 körül, Telekösszeírás, V. füzet: Nagymező utca 13.
1850 körül-1862 körül, Telekösszeírás, IV., VI. füzet: 1187
1835-1847 körül, Telekösszeírások, II., III. füzet: 1112
1824 körül, Telekösszeírás, I. füzet: 1003
1824 előtti telekösszeírási szám: 804
1786, József-kori telekkönyv mutatója: 590
Grundbuch, Telekkönyv: 1166




Budapesti levéltári iratok a térképen

29337. helyrajzi szám kapcsolódó levéltári iratai

Előző: 29336 | Következő: 4555/65

Adatbázis megnevezése	Kapcsolódó rekordok
Telekkönyvi betétek adatbázisa BFL	HU BFL - XV.37.c - 3953
Lakás adatszolgáltatási ívek, 1944 BFL	Nincsenek kapcsolódó rekordok
Építészeti tervek BFL	Kapcsolódó levéltári iratok megtekintése

Fotók, képeslapok a környékről



1. ábra. A 29337 helyrajzi számú telek adatlapja (a telek helye a térképen, története, kapcsolódó levéltári iratok)

Ezek elsősorban katonai célból készültek, így a háromdimenziós, a domborzatot jól érzékeltető ábrázolási módnak is teret adtak.

Az Osztrák Császárság által 1760-ban kezdett ún. első katonai felmérés (*Josephinische Landesaufnahme*) alkalmával a teljes birodalmat felmérték, és kézzel festett, egy példányos térképek készültek 1:28800-as léptékkel. Ez a korabeli hadászatnak nyújtott felbecsülhetetlen segítséget, hisz részletessége okán esetleg előfordulhat, hogy az összes falu összes háza vagy telke életszerűen és pontosan ábrázolt. Az ekkor alkalmazott térképészeti módszerek azonban csak kb. 500–600 méteres pontosságot tettek lehetővé. Ezzel szemben a második katonai felméréssel (*Franziseische Landesaufnahme*, 1806) már olyan térképek készültek a császárságban, ahol 100–150 méteres a hibátűrés, illetve a jelenlegi elérő a georeferálás pontossága. A párhuzamosan indított (1817) kataszteri felmérés, amely ugyancsak a teljes birodalomra kiterjedt (és 1900-ig tartott) 1:2880-as léptékével soha nem látott részletességet és pontosságot eredményezett. Itt akár 15–20 méteres pontosságot is sikerült elérniük a térképészeknek.

A fenti alapokon indult el Budapest (1873 előtt Pest, Buda és Óbuda) térképezése, amelynek eredményeit ma legteljesebben Budapest Főváros Levéltára (BFL) őrzi. A digitalizálás és a feldolgozás ezen térképek alapján történt és történik folyamatosan. A térképek feldolgozásának első lépése a térképlapok szkennelése (300 dpi, true-colour, tömörítetlen JPG), amit az említett georeferálás követ. A georeferálás során fel kell tárnai az eredeti térképek készítői módját, így

- meg kell állapítani a térkép vetületi rendszerét (hiszen a térkép a háromdimenziós Föld két dimenzióba vetített képe),
- hasonlóan a vetületi paramétereket (a vetítés középpontja, a szelvényezés módja),
- a térképkészítők által feltételezett ellipszoidot és annak középpontját (hiszen minden vetítésnél ezek a paraméterek a meghatározók).⁷

Térképrétegek egymáson

A szakirodalomból tudjuk, hogy a budapesti térképezésnél a 19. és 20. században az ún. Gellérthegy középpontú sztereografikus vetületi rendszert alkalmazták, melynek paraméterei jól ismertek. Így a feladat az, hogy az alkalmazott GIS-rendszernek⁸ megadjuk ezeket a paramétereket, majd a térképek négy sarkát kézzel (esetleg automatikus módon) megjelöljük, ezután a program minden egyes pixelre kiszámolja az adott pont GPS-koordinátáit. Így a szelvények összeilleszthetők és a sok száz (vagy akár több ezer) szelvényből egyetlen nagy, összefüggő térkép készülhet. Ha ezt szokásosan a Google által használt Mercator-vetületre konvertálják, az is lehetővé válik, hogy a térképeket bármely mai (Google Maps, Google Satellite, OpenStreetMap stb.) térképre rávetítsék. A térképek publikálásánál a sokak által használt OpenLayer nyílt forráskódú szoftvert használták a szakemberek.⁹

A *Budapest Időgép* további alapját a Budapest Főváros Levéltárában végzett nagy tömegű digitalizálások teremtették meg: jelen esetben a telekkönyvi betétek, építészeti tervek és a gazdag fotóanyag. Az iratdigitalizálás számos más forrástípust is felölel – mint később látni fogjuk – a fejlesztés első körében azok kerültek a szolgáltatásba, amelyek adatbázisrekordjai eleve tartalmazták az érintett telek mai helyrajzi számát.

A fent vázolt munkamenetnek természetesen vannak egyéb, nagyon fontos részfeladatai is. A *Budapest Időgép* megvalósításának egyik első és nagyon fontos kihívása az volt, hogy a szkennelt és georeferált térképeket vektorizálni kellett. Ennek során berajzolták az utcákat, tereket és a térképen ábrázolt és névvel ellátott (ebből kifelé fontosnak tartott) épületeket. Így időszakonként (térképenként) előállították az adott időszak utcaneveit (a hozzájuk tartozó vonallal), a városrészeket, földrajzi objektumokat (hegyek, völgyek, csúcsok), olyanokat is, amelyeknek területe nem meghatározható, valamint a névvel jelzett épületeket. Az eredmény egy kereshető adatbázis lett, amelynek segítségével régi utcanevekre, intézményekre (épületekre) is kereshetünk. A vállalkozás során az alábbi térképeket dolgozták fel:

⁷ Talán érdemes itt megemlíteni, hogy a Föld pontos alakját és középpontját csak 1984-ben sikerült pontosan megállapítani, ez a ma használt WGS84 ellipszoid.

⁸ A GIS, az angol *Geographic Information System* rövidítése, mely egy földrajzi információs rendszer, elterjedtebb néven térinformatikai vagy geoinformatikai rendszer, ez esetben ennek egy konkrét implementációja.

⁹ A georeferálás az ELTE TTK Geofizikai és Űrtudományi Tanszék munkatársai által kifejlesztett módszer szerint történik; itt érdemes név szerint is megemlíteni Tímár Gábort és az informatikai fejlesztők vezetőjét, Biszak Elődöt.

- Pest-Buda–Óbuda áttekintő térképe a jelentős középületek rajzával és látképekkel (Vasquez), 1837;¹⁰
- Pest és Buda kataszteri térképsorozata az 1872–1920 közötti változások utólagos jelölésével, 1867–1872/1873;¹¹
- Budapest közigazgatási térképsorozata, 1908;¹²
- Budapest közigazgatási térképsorozata az épületek szintszámával és az újabb utcanevekkel, 1937.¹³

Mit és hogy kereshetünk?

Egy keresőkérdés megadása után mind a négy térképen egyszerre indul a keresés, az eredmény pedig jelzi, hogy melyik térképen van a találat, és az (pl. utcavektor vagy épületkörvonal) kivilágításra is kerül. Akár iskolákra, akár kórházakra is kereshetünk, de persze a legfontosabb eredmény, hogy szinte tetszőleges időpontú utcanevekre is visszakereshetünk. Így akár a 19., akár a 20. századból származó a forrás, amely utcanevet említ, könnyedén beazonosítható az utca helye azonos korú térképen.¹⁴ Ugyanígy járhatunk el intézmények, épületek, de akár városrészek vagy általánosabb földrajzi helyek esetében is. A vektorizált térképek egymásba áttűntethetők mint egymásra rakott rétegek. Ennek legnagyobb előnye – a látványos hatáson túl –, hogy akár a teljes város, akár egy-egy részletének változásai vizuálisan is követhetők, illetve a konkrét keresés eredményét láthatjuk különböző korokba helyezve is. Így például könnyedén azonosítható egy mai térképen, hol élt egy adott személy, hol történt egy fontos esemény, mi a mai utcaneve, házszáma egy régi történeti lokalizációnak.

Levéltári adatok a térképeken

A *Budapest Időgép* nevű alkalmazás másik – s igen lényeges – része a levéltári és egyéb adatok ábrázolása a térképre vetítve. Ehhez első körben a levéltári adatok *geokódolása* szükséges. Ennek keretében a lakcím (régi vagy mai), illetve a helyrajzi szám alapján meghatározzák az adott irat vonatkozásainak GPS-koordinátáját. A feladat végrehajtásának elengedhetetlen része volt a helyrajzi számok adatbázisának elkészítése, amely a GPS-koordinátát is tartalmazza.

¹⁰ Pest-Buda–Óbuda áttekintő térképe, 1837, hozzáférés: 2018.05.22, <http://mapire.eu/hu/synchron/budapest/?left-layers=osm%2C26&right-layers=osm%2C26&bbox=2101463.40790064%2C6001392.954991487%2C2151529.66142743%2C6044350.564887756>.

¹¹ Pest és Buda kataszteri térképsorozata, 1867–1872/1873, hozzáférés: 2018.05.22, <http://mapire.eu/hu/synchron/budapest/?left-layers=osm%2C26&right-layers=osm%2C26&bbox=2101463.40790064%2C6001392.954991487%2C2151529.66142743%2C6044350.564887756>.

¹² Budapest közigazgatási térképsorozata, 1908, hozzáférés: 2018.05.22, <http://mapire.eu/hu/synchron/budapest/?left-layers=osm%2C26&right-layers=osm%2C26&bbox=2101463.40790064%2C6001392.954991487%2C2151529.66142743%2C6044350.564887756>.

¹³ Budapest közigazgatási térképsorozata, 1937, hozzáférés: 2018.05.22, <http://mapire.eu/hu/synchron/budapest/?left-layers=osm%2C26&right-layers=osm%2C26&bbox=2101463.40790064%2C6001392.954991487%2C2151529.66142743%2C6044350.564887756>.

¹⁴ Természetesen ez nem pótolja egy digitális utcanév-lexikon meglétét, egy olyanét, amelyben az utcanevek keletkezésének és módosulásának időpontjai is megbízhatóan fellelhetők.



2. ábra. Az átlátszóság állításával a régi és az új térképek összevethetők

Jelenleg az alábbi adatok térképi ábrázolása látható:

- telekkönyvi betétek,
- lakásívek,
- tervek,
- topográfia.

Néhány szót érdemes mondani a kevésbé ismert *telekkönyvi betétekről*.¹⁵ A fővárosi ingatlanokra vonatkozó, a modern telekkönyvi rendszer alapját jelentő telekkönyvi betétek Budapest nagyvárossá fejlődésének időszakából hitelesen és teljes körűen tartalmazzák az ingatlanok állagában, tulajdonviszonyaiban beállott változásokat, valamint a rájuk betáblázott terheket. Így önmagukban is sokat elárulnak a városépítés gazdasági és társadalmi háttéréről, de mivel az építkezésekre vonatkozó hivatali iratokat általában a korabeli tulajdonos nevéből kiindulva lehet megtalálni, nélkülözhetetlen alapjai mindennemű építészettörténeti, helytörténeti és az ingatlanokat egyéb szempontból érintő kutatásnak, továbbá napjainkig ható ún. jogbiztosító erővel is rendelkeznek. A 1870-es évektől felvett – a budai oldalon az 1850-es évekig visszanyúló – betétek vezetését a budai oldalon 1921-ben, a pesti oldalon 1926-ban szüntették meg. Ekkor vezették át az ingatlanok adatait az immár a jelenleg is élő helyrajzi számozási rendszerben készült új betétekbe, amelyek az 1970-es évekig, az egységes ingatlan-nyilvántartási rendszer kialakításáig voltak használatban, és a ma érvényes tulajdoni lapok közvetlen előzményét képezik.

Jelenleg a „Pesti telekkönyvi betétek” (BFL XV.37.c) állagából az 1–15000. számú betét, valamint a fennmaradt összes óbudai betét (BFL XV.37.b) digitális másolata kereshető a *Budapest Időgép* portálon. A georeferált térképen navigálva és a megfelelő telekre kattintva az adott helyrajzi számokhoz tartozó betétek jelennek meg.

¹⁵ Ezt az 1886. évi XXIX. törvénycikk szabályozta részletesen „a telekkönyvi betétek szerkesztéséről” címmel.

A lakásíveknek külön történeti jelentőségük van, elsősorban a vészkorszakra vonatkozóan. Mint ismeretes, a zsidónak minősített budapesti lakosságot 1944 júniusában sárga csillaggal megjelölt házakba költöztették össze. Ennek előkészítése gyanánt Budapest helyettes polgármestere 1944. május 30-án elrendelte valamennyi fővárosi lakás és bérlő összeírását. Az összeíráshoz kiadott és június 1-én szétszított adatszolgáltatási íveket a háztulajdonos vagy házfelügyelője 24 órán belül köteles volt kitölteni. A formanyomtatványokon a lakások lakbérét, szobaszámát, utcai vagy udvari fekvését, valamint a háztulajdonos és valamennyi bérlő „zsidó” vagy „nem zsidó” mivoltát kellett feltüntetni.¹⁶

Mindeddig úgy tudtuk, hogy a kitöltött adatszolgáltatási ívek nem maradtak fenn. Feltételeztük, hogy – miként a székesfőváros lakásügyi ügyosztályának teljes második világháború alatti anyaga – ezek is megsemmisültek az ostrom alatt. 2015 szeptemberében azonban az V. kerület, Kossuth tér 13–15. számú ház egyik lakásának felújítása során a közfal kibontásakor mintegy 7000 oldalnyi ív került elő. Az adatbázis így teljes terjedelemben tartalmazza a digitalizált íveket, a rajtuk található minden érdemi adat kereshető. A rekordok tartalmazzák a lakóházak adatait: cím, helyrajzi szám, az épület magassága, építési éve, a zsidónak minősített (sárga csillag viselésére kötelezett), valamint az erre nem kötelezett (nem zsidónak minősített) tulajdonosok és főbérlők számát, valamint a lakás minősítését (zsidók vagy nem zsidók által lakható). A tulajdonosok nevét és lakcímét, valamint a főbérlők nevét, lakásának emelet- és ajtószámát, a lakás éves bérét, szobaszámát, utcai vagy udvari fekvését.

Az építészeti tervek a legkülönbébb fondokból és gyűjteményekből jönnek össze, mint például a budai és pesti építőbizottmány adatai, Ybl-tervek stb. Hangsúlyozni kell, hogy egyelőre főként az 1873-as városegyesítés előtti tervek, illetve az említett Ybl-hagyaték tervei érhetők el. Az ez utáni korszak tervei egyelőre viszonylag kis mennyiségben kerültek feldolgozásra, így például a Buda belső részéről szólóak. Ezek tömegesebb megjelenése a rendszerben az idei évre tervezett fejlesztésektől remélhető.

A rendszer használatakor a felhasználó által meghatározott térképen (a fent említett négy közül választhatunk, illetve ezek vektorizált változatát is használhatjuk, de természetesen a mai térképet is kezelhetjük) jeleníthetjük meg vagy az összes adatot, vagy csak az általunk kiválasztott adatokat (pl. a telekkönyvi adatokat). Ekkor a térképen az adott területen található helyrajzi számmal megjelölt és adatokkal ellátott telkek számát látjuk. Ha közelítünk, egyre kisebb területen egyre kisebb számot látunk, majd elérhetjük a helyrajzi szám szintjét is. Az adatot jelentő körre kattintva megjelenik az adott helyrajzi számú rekord, ahol jelezve van, hogy milyen adat tartozik az adott telekhez. Mindig megkapjuk az adott telek topográfiai adatait (a hetvenes évek óta megjelenő topográfia alapján), kattinthatunk azon a levéltári iratra, amely az adott telekhez kapcsolódik, és ott megtekinthetjük magát a digitalizált iratot is (terveket, telekkönyvi betéteket, lakásíveket).

¹⁶ Minderről részletesebben lásd „»Ki kell menni a zsidónak a lakásából is«: A csillagos házak kijelölése Budapesten,” Budapest Főváros Levéltára, hozzáférés: 2018.05.22, <http://holocaust.archivportal.hu/temak/ki-kell-menni-zsidonak-lakasabol-csillagos-hazak-kijelolese-budapesten>.

Képek és térképek

A *Budapest Időgép* érdekessége, hogy mindig megkapjuk a helyről vagy környékéről származó fotókat vagy képeslapokat is. Ennek alapja tulajdonképpen a régebbi adatbázis és képtároló szolgáltatások integrációjában rejlik. A Hungaricana-adatbázisok „Képcsarnok” fejezetében rengeteg fotógyűjtemény található. Most csak a – témánk szempontjából – legfontosabbakat említjük meg:

- A szerencsi Zempléni Múzeum képeslap-gyűjteménye (mintegy 200000 városkép)
- FSZEK Budapest Gyűjtemény fotói,
- BFL Klösz-fotói,
- BFL a Fővárosi Tanács VB Városrendezési és Építészeti Osztályának fotói (ez lényegében az összes főút valamennyi házát ábrázolja az ötvenes-hatvanas évekből),
- Fortepan-gyűjtemény.

Valamennyi fotó GPS-koordinátával ellátott, így az egy adott hely környezetében lévő fotókat könnyedén le tudjuk kérdezni, tehát az adott telek körül meg tudjuk mutatni a különböző időből származó felvételeket.

Érdekesség, hogy a *Budapest Időgép* oldaláról is elérhetők a Budapestről 1944-ben készült légi felvételek. A Magyar Királyi Honvéd Légierő 1944 áprilisában, az első nagyobb angolszász bombázások után pár nappal, kis magasságból végigfényképezte Budapest találatokat kapott kerületeit. A nagyjából 550 légi felvétel negatívjai a közelmúltig csak a Hadtörténeti Levéltárban voltak kutathatók. Ezek az anyagok váltak itt elérhetővé az Arcanum és a Fortepan közös erőfeszítéseinek köszönhetően, mégpedig precíziós összeillesztéssel, a feleslegek kivágásával, a torzítások csökkentésével és természetesen georeferálással ellátva.

A továbbfejlesztés irányai, a felhasználás lehetőségei

A nemrég publikált *Budapest Időgép*-alkalmazás gyorsan elnyerte a felhasználók tetszését, nagyon sokan a kutatás egy fontos eszközének tekintik. Mindez ösztönzést adott a rendszer továbbfejlesztésére. Ez egyrészt az adatok számának növelését (több építészeti terv, több telekkönyvi betét), másrészt az eltérő adatok bevonását jelenti. Ennek a munkának a keretében folyik egy adott időszak (1916) utca-házzám adatbázisának előállítása, a hozzátartozó lakástulajdonosok adatainak feltüntetésével.¹⁷ Emellett érdemes nemcsak a térképről leolvasni az intézmények adatait (hiszen nagyon sok nem jelenik meg itt), hanem a *Budapesti Lakcímjegyzékből* kiemelni például az összes iskolát, összes kórházat, összes templomot, összes éttermet, kávéházat, majd azt ábrázolni a térképen (a létrehozott utca-házzám adatbázis felhasználásával).

Ugyancsak fontos, hogy ne csak az 1920 utáni nagy átszámozást követő (mai) helyrajzi számokról legyen adatbázisunk, hanem az ezt megelőző időszakról is (hiszen

¹⁷ Akkoriban általában bérelték a lakásokat; a teljes főbérleti lista feldolgozása az adatbázisban egyelőre nem történik meg.

lehet, hogy egy forrásban az 1901-es helyrajzi számot találjuk meg). Így hamarosan elkészül egy, az 1920-as éveket megelőző helyrajzi számokat tartalmazó adatbázis is. Az adatok körét pedig az 1914–1918 közötti közjegyzői ügyek, iratok térképre helyezésével bővítik, így egzakt képet kapunk majd arról, hogy milyen helyekről, mely városrészekből származnak az iratok, azoknak milyen a megoszlása. Továbbá egy adott házra kattintva ideális esetben a tulajdonváltások, az építészeti átalakítások, az ott működő cégek, kávéházak stb. adatait együtt láthatjuk az érintettekhez (tulajdonos, lakó, kereskedő stb.) köthető közjegyzői ügyekkel.

A lehetséges hasznosítási területekről jól informálnak az egyes kutatási aspektusok. Ilyen például az 1956-os forradalom és szabadságharc áldozatainak térinformatikai adatbázisa. Ez tartalmazza az 1956. október 23-tól 1957 tavaszáig erőszakos halállal, nagy valószínűséggel a forradalmi eseményekkel összefüggésben elhunyt személyek adatait, és azokat meg is jeleníti, illetve kereshetővé teszi a térképeken. Ennek alapja, hogy a budapesti halotti anyakönyvekben szereplő információkat más rendelkezésre álló forrásokkal egybevetve felderíthető volt a halottak száma, mely 2500–3000 főre tehető. A térinformatikai adatbázisban jelenleg a Budapesten elhunyt körülbelül 1700 nevesíthető áldozat adatai szerepelnek mai és közel egykorú térképre vetítve, kereshető formában. A térképen megjelenő körök nagysága és a bennük szereplő szám az adott városrészben a forradalom és szabadságharc küzdelmeiben vagy azokkal összefüggésben elhunyt áldozatok számát jelzi. A térképre nagyítva a körök a halálesetek helyszínei szerint elkülönülnek. Ha a kurzort a körök fölé visszük, megjelennek az adott helyszínhez köthető áldozatok halálára vonatkozó főbb adatok. A számozott körökre kattintva az áldozatok személyes adatai és a halálesetre vonatkozó részletes információk is láthatóvá válnak. Emellett lehetőség van a keresés szűkítésére, az adatok időpont és név szerinti szűrésére is.¹⁸

Nagyon fontos látni azt, hogy az adatok és térképek ilyen szintű és mennyiségű feldolgozása és összekapcsolása az, ami valódi hozzáadott értéket hoz létre. Valójában a felhasználónak rengeteg lehetséges lekérdezése és ebből fakadó következtetése adódhat. A levéltári anyag feldolgozása, majd százezernyi adat összekapcsolása térképi pontokkal, az utcák, épületek megtalálása stb. maga is tudományos munka. Nyugodtan mondhatjuk, hogy egy ilyen magas szinten fejlett eszköz önmagában digitális bölcsészeti projekt, de bármilyen más rendszerrel, adatforrással összekapcsolva, korreláltatva és lekérdezve sok más testvértudomány forrásává válhat. A térképszelvények időbeni síkjainak egymásra vetítése és átjárhatósága, a térképpontoknak földrajzi aspektussal is rendelkező adatokkal való összekötése tehát számos lehetőséget rejt. Ez a fajta technológia nagy mennyiségű forrásadattal (*big data*) történő megvalósítása nemzetközi szinten is ritka, így a projekt módszertani jelentőséggel is bír.¹⁹

¹⁸ Vö. Tulipán Éva és Molnár Gábor, „Az 1956-os forradalom és szabadságharc budapesti halottainak térinformatikai adatbázisa,” *Betekintő* 3. sz. (2016), http://www.betekinto.hu/2016_3_tulipana_molnar.

¹⁹ Hasonló térképes-adatelemzéses online projektről alig pár darabról tudunk a világban. Magyar nyelven az egyik ilyen projektről bővebben lásd: Horváth Bence, „Időgéppel tárják fel Velence múltját,” *444*, 2017. júl. 12., <https://444.hu/2017/07/12/idogeppel-tarjak-fel-velence-multjat>. A létező néhány kezdeményezés összefogásával egy nagyobb szabású EU-projektre is indult iniciatíva: Time Machine FAT Flagship, hozzáférés: 2018.05.25, <http://timemachineproj>

Mondhatjuk, hogy a történelem tanulmányozásának egy olyan új eszközét kapjuk kézbe, amellyel térben és időben szemlétesebben, nemritkán új ismeretekhez jutva járhatjuk be a múlt geográfiai vetületeit.

Budapest Time Machine

The *Budapest Time Machine* (*Budapest Időgép*, <https://hungaricana.hu/hu/budapest-idogep/>) is a recent project that offers innovative ways to use historical maps as a research tool. The web service contains overlapping maps from different periods with attached, geospatial, archival materials and photographs. These data present the recent history of Budapest and its geographic dimensions. This study demonstrates novel ways in which this resource can be used for research purposes.

Keywords:

historical geography, digital maps, georeferencing, time machine, Budapest Time Machine

ect.eu/. Az ebben részt vevő időgép-projektek közül egyébként a budapesti az egyetlen, amely már élesben működik, mégpedig nagy adatmennyiséggel feltöltve.

Ruttkay Zsófia

Moholy-Nagy Művészeti Egyetem TechLab

ruttkay@mome.hu

Digitális Múzeum – a MOME TechLab projektjeinek tükrében

A tanulmány tárgya a digitális technológiák múzeumi alkalmazása. Bevezetőként körbejárja a digitális múzeum terminológiai értelmezésének alternatíváit. A szerző arra fókuszál, hogy a digitális technológiák alkalmazásával miként bővíthető ki a múzeumlátogatás élménye időben és térben, valamint miként motiválható a látogató aktív részvételre. Tizenöt projekt kerül bemutatásra, melyek oktatási környezetben hallgatói projektként vagy múzeumi megrendelésre, a TechLab Fellows alkotói kollektíva által készültek. Minden esetben a partner kulturális örökség intézménnyel szoros együttműködésben születtek meg az alkalmazások, melyek tartalma változatos, az irodalomtól a zenén át a képzőművészetig. A projekteket a kiállítási funkciójuk és mechanizmusok szerinti csoportosításban mutatjuk be, olyan fogalmak mentén, mint például a meglepetés és az érzelmek keltése, játékos felfedezés, interaktív adatvizualizáció, műtárgyak értelmezése plusz információk csatolásával, a látogató megszólítása, illetve virtuális kiállítás-látogatás a falakon belül és kívül. A legtöbb projektről demo anyag is elérhető a <http://techlab.mome.hu/> oldalon. A tanulmány a digitális technológiák sikeres alkalmazásához elengedhetetlennek tartja az interdiszciplináris együttműködést, valamint az eredményesség kritériumainak tudományos igényű vizsgálatát.

Kulcsszavak:

digitális múzeum, virtuális kiállítás, MOME TechLab



1. Bevezető: a *digitális múzeum* meghatározása

A *digitális múzeum* kifejezés érdekes módon korábban tűnt fel, mint a *digitális bölcsészet* (*digital humanities*).¹ Ben Davis 1994-ben az MIT Museum egy kiállítása kapcsán született cikkében széles kontextusban, a kurátorközpontúból a digitális technológiák kiaknázása révén látogatóközpontúvá váló múzeumot nevezi digitális múzeumnak felsorolva a technológia bevetésének különböző (elvi) lehetőségeit és rögtön dilemmáit is.² A víziószerű felvetéseket aztán utolérte a technológia fejlődése, és az évek során

¹ Matthew G. Kirschenbaum, „What is Digital Humanities and What’s It Doing in English Departments?” *ADE Bulletin* 150. sz. (2010): 55–61.

² Ben Davis, „Digital Museums,” *Aperture* 136. sz. (1994): 68–70.

egyre többféle speciális jelentéssel – a múzeumok webes megjelenésével,³ az állománydigitalizálással, a virtuális múzeummal, majd a kiállítási interpretációs technológiákkal összefüggésben⁴ – használták a terminust. Ma már külföldön igen elterjedt a digitális technológiák alkalmazása, jó gyakorlatokról, tanulságokról, sőt múzeumok által kezdeményezett kutatásokról is lehet olvasni.⁵

A digitális múzeum terminust 2010-ben kezdtem használni egy egyetemi kurzus címeént.⁶ Itthon tudtommal elsőként egy PIM-ben publikált tanulmány címében szerepelt a kifejezés,⁷ melyben a szerző elsősorban a gyűjtemény digitalizálásának és egységes múzeumi elektronikus szolgáltató környezet létrehozásának a szükségességét és mikéntjét tárgyalja, de egy további célként megjelöli a „fizikailag megépített kiállítások informatikai támogatását” is. Noha a digitális bölcsészet saját folyóiratokkal, könyvekkel, egyetemi kutató-, oktatóközpontokkal és projektek sokaságával büszkélkedhet,⁸ máig eltérnek a nézetek arról, hogy mi minden tartozik tevékenységi körébe a szövegek számítógépes eszközökkel történő kutatásán túl.⁹ A *digitális múzeum* kifejezés kevésbé elterjedt, ezért bevezetésként tisztázom, mit is értek alatta. Az interneten és elméleti vagy éppen gyakorlatorientált publikációkban a digitális múzeum alábbi, specifikus, szűk jelentéseire bukkanhatunk, ahol esetenként a *virtuális* jelző fordul elő a szókapcsolatban:

³ J. Takahashi, T. Kushida, Y-K. Hong, S. Sugita, Y. Kurita, R. Rieger, W. Martin, G. Gay, J. Reeve and R. Loverance, „Global Digital Museum: Multimedia Information Access and Creation on the Internet,” in *Proceedings of the Third ACM Conference on Digital Libraries* (New York: ACM, 1998), 244–253, <https://dl.acm.org/citation.cfm?doid=276675.276703>.

⁴ George F. MacDonald and Stephen Alford, „Conclusion: Toward the Meta-Museum,” in *The Wired Museum: Emerging Technology and Changing Paradigms*, ed. K. Jones-Garmil (Washington: American Association of Museums, 1997), 268–278; Ross Parry, ed., *Museums in a Digital Age* (London & New York: Routledge, 2009).

⁵ D. Theuns, C. Verweij, B. Schrandt, R. Van Westen and W. Van den Heuvel, „Do (Not) Try This at the Museum: Lessons from Digital Museum Lab,” *Museum ID*, 2018. jan. 3., <http://museum-id.com/not-try-museum-lessons-learned-digital-museum-lab/>.

⁶ Bényei Judit and Ruttkay Zsófia, „Digital Museum: A Multidisciplinary University Course,” előadás a *Perspectives on Art Education – Symposium – Conversations Across Cultures* c. konferencián, University of Applied Arts Vienna, 2015. május 28–30.

⁷ Bánki Zsolt, „Digitális múzeum: Egységes múzeumi elektronikus szolgáltatási környezet kialakítása a Petőfi Irodalmi Múzeumban,” *Tudományos és Műszaki Tájékoztatás* 54, 9. sz. (2007), http://tmt-archive.omikk.bme.hu/show_news.html?id=4776&issue_id=486.html.

⁸ „CUNY Academic Commons Wiki Archive,” hozzáférés: 2018.05.22, https://wiki.common.gc.cuny.edu/Sample_Projects/.

⁹ Patrik Svensson, „The Landscape of Digital Humanities,” *Digital Humanities Quarterly* 4, 1. sz. (2010), <http://digitalhumanities.org:8081/dhq/vol/4/1/000080/000080.html>; Niels Brügger, „Digital Humanities in the 21st Century: Digital Material as a Driving Force,” *Digital Humanities Quarterly* 10, 2. sz. (2016), <http://digitalhumanities.org:8081/dhq/vol/10/3/000256/000256.html>.

- Egy, csak virtuálisan, az interneten létező múzeum.¹⁰
- Egy valódi múzeum vagy egy-egy kiállítás online lenyomata, prezentációja, sokszor a múzeum vagy kiállítás térbeli bejárását is lehetővé téve (gyakran a *virtual museum* névvel különböztetve meg ezt a hagyományos múzeumtól).¹¹
- Létező múzeumok tematikus¹² vagy földrajzi/történeti csoportjának interneten megvalósított *metaintézménye*, ahonnan az egyes múzeumok gyűjteményei közzösen, online érhetők el.¹³
- Egy olyan kiállítás, amelyben nincsenek jelen műtárgyak fizikai valójukban, csak digitális másolat vagy rekonstrukció formájában,¹⁴ a digitális interpretációs eszköztárat – például szimuláció, virtuális valóság – felvonultató magyarázatokkal.
- Olyan, műtárgyak mozgatása nélkül utaztatható, fizikai helyhez kötött kiállítás, ahol egy hagyományos múzeumot és műtárgyakat VR-szemüveggel lehet felidézni és bejárni¹⁵ – ebben az esetben a fizikai hely nem múzeumépület, hanem a VR-környezet működtetéséhez szükséges terem.
- Olyan időszaki kiállítás vagy egy témának szentelt múzeum a virtuális térben, melynek tárgyai és egyéb emlékei¹⁶ a világon szétszórva találhatók meg több intézményben. Ezek digitális másolata – a valós kiállítások létrehozásakor megszokott kutatás és kurátori munka eredményeként – kerül egy missziót, üzenetet közvetítő narratív egységbe a virtuális felületen, esetleg kisebb fizikai utazórészleggel kiegészülve.¹⁷
- Eleve digitális formában létrehozott műtárgyak fizikai¹⁸ vagy virtuális¹⁹ térben bemutatott gyűjteménye.

A fenti, teljességre törekvő felsorolás azt körvonalazza, hogy mi minden esik *kívül* a digitális múzeum általam is értett és legújabbán néhány publikációban,²⁰ valamint

¹⁰ Virtual Shoe Museum, hozzáférés: 2018.05.22, <http://www.virtualshoemuseum.com/>.

¹¹ Smithsonian National Museum of Natural History, hozzáférés: 2018.05.22, http://naturalhistory.si.edu/VT3/NMNH/z_NMNH-016.html.

¹² „Digital Museum,” Natural History Museum, hozzáférés: 2018.05.22, <http://www.nhm.ac.uk/our-science/our-work/digital-museum.html>.

¹³ Digitalt Museum, hozzáférés: 2018.05.22, <https://digitaltmuseum.no/>.

¹⁴ Xinhua, „3D Digital Exhibition Brings Sensational Archeological Finding from China to Singapore,” *People’s Daily Online*, 2012. okt. 11., <http://en.people.cn/202936/7973662.html>.

¹⁵ Molly Schuetz, „Virtual Reality Museum Puts Rembrandt in High-Def,” *Bloomberg*, 2017. okt. 26., <https://www.bloomberg.com/news/articles/2017-10-26/virtual-reality-museum-puts-rembrandt-in-high-def>.

¹⁶ Virtual Museum, Museo Galileo, hozzáférés: 2018.05.22, https://catalogue.museogalileo.it/?_ga=2.9053239.320339754.1511974246-1175795331.1511974246.

¹⁷ Digital Museum of Global Buddhist Cultural Heritage-projekt, Museum of Archaeology and Anthropology, hozzáférés: 2018.05.22, <http://maa.cam.ac.uk/digital-museum-of-global-buddhist-cultural-heritage/>.

¹⁸ Dutch Digital Art Museum Almere, hozzáférés: 2018.05.22, <https://ddama.eu/>.

¹⁹ Digital Art Museum, hozzáférés: 2018.05.22, <http://dam.org/>.

²⁰ Ali Hossaini and Ngaire Blankenberg, eds., *Manual of Digital Museum Planning* (Rowman & Littlefield Publishers, 2017).

nemzetközi fórumon is tárgyalt értelmezésén.²¹ Ezen felfogás szerint a digitális múzeum alatt értendő minden olyan, a digitális technológiák tárházát kiaknázó eszköz, módszer, mely fizikailag létező gyűjteménnyel és azok szakszerű megőrzésével, tudományos igényű feldolgozásával és szélesen értendő oktatási célú bemutatásával foglalkozó kulturális örökség intézmények – hagyományos múzeumok, levéltárak – munkáját segíti és terjeszti ki.

Eszköz alatt könnyen használható programozási keretrendszereket, egy-egy feladatra kifejlesztett modulokat (például képfeldolgozás, dinamikus mozgásszimuláció, kibővített valóság), illetve speciális, egyedi célú alkalmazásokat is értünk, melyek használata persze nem képzelhető el egy vagy több digitális fizikai berendezés (processzor, szenzorok, kamera, illetve ilyeneket magukba foglaló tablet vagy okostelefon) nélkül.²²

A múzeumok világában a digitalizált archívum, a digitális másolatok (fotók, 3D-modellek) készítése, illetve eleve digitális formában gyűjtött anyagok (például hangfelvételek) tárolása, a nyilvántartás formai és tartalmi jegyei, a restaurálásban és kutatásban használatos eljárások (például képalkotó technikák, rekonstrukciót és beazonosítást segítő számítógépes grafikai- és mesterségesintelligencia-módszerek) több évtizedes múltra nyúlnak vissza, ahogy a támogatott tevékenységek is hosszú ideje a múzeum elfogadott alapfeladatai.

A prezentáció, interpretáció és tágabb értelemben a múzeum részéről a társadalom széles körei számára nyújtandó szolgáltatás azonban csak évtizedek alatt vált tevékenységi körük elfogadott, sőt kiemelt részévé.²³ Ráadásul e szolgáltatások mibenléte ma, az információs társadalom és globalizáció korában jóval összetettebb kérdés, mint azelőtt, amikor a művelődni vágyó közönség zarándoklatszerűen, az ismeretszerzés igényével nézett végig vitrinsorokban felhalmozott tárgyakat, falakra akasztott festményeket, betartva a „múzeum mint a tudás szentélye” látogatásának protokollját.

A múzeum mint közszolgálati intézmény megújulása jóval a digitalizáció és az internet megjelenése előtt megindult a társadalom szélesebb rétegeinek megszólítása, a múzeum mint intézmény demokratizálódása érdekében.²⁴ A gazdag szakirodalom²⁵ és a vezető múzeumok példaértékű praxisa – célzott tárlatvezetések, előadások, audio guide-ok, gyerekeknek készült kiadványok és programok, a múzeumi gyűjtemény elemeinek megjelenése a falakon kívül, például a boltban beszerezhető hétköznapi tárgyakon – ellenére az új szerepkör, az élethosszig tartó tanulás megvalósítása még ma

²¹ „The Digital Museum,” Tumblr, hozzáférés: 2018.05.22, <https://thedigitalmuseum.tumblr.com/>; Museum Digital Transformation, hozzáférés: 2018.05.22, <http://museumdigitaltransformation.it/>.

²² Lásd Jasper Visser, „Strategies for a Heritage Revival in the Digital Age,” in *Companion to European Heritage Revivals*, eds. Linde Egberts and Koos Bosma (Springer International Publishing, 2014), 73–93, https://doi.org/10.1007/978-3-319-07770-3_4; Bényei Judit és Ruttkay Zsófia, „Kulturális örökség közvetítése digitális, interaktív technológiák segítségével,” *Szociálpedagógia* 1–2. sz. (2015): 36–52.

²³ Gail Anderson, ed., *Reinventing the Museum: The Evolving Conversation on the Paradigm Shift*, 2. kiadás (Lanham, Md.: AltaMira Press, 2012).

²⁴ Az ICOM időről időre felülvizsgálja a múzeum definícióját. Még a 2007-es meghatározás van életben, de már dolgoznak a frissítésén. „Museum Definition,” International Council for Museums, hozzáférés: 2018.05.22, <http://icom.museum/the-vision/museum-definition/>.

²⁵ Anderson, *Reinventing the Museum*, 191–286; Frazon Zsófia, *Múzeum és kiállítás: Az újrarajzolás terei* (Budapest–Pécs: Gondolat–PTE Kommunikáció- és Médiatudományi Tanszék, 2011), 177–185.

is a kísérletezés terepe, ami a műfaji határokat, az eszközöket és azok eredményességét illeti.

Az új társadalmi küldetés megvalósítása az információs társadalom, a digitális bennszülöttek korában újabb kérdéseket vet fel mind a potenciális közönség elérését, mind az ismeretátadás módjait illetően.²⁶ Egyrészt utat kell találni a mai fiatalokhoz, alkalmazkodva a megváltozott információszerzési, kommunikációs és szabadidős szokásaikhoz, hogy ne szürküljön el a múzeumlátogató közönség, és a következő generációk számára releváns, általuk is fenntartandó intézmény legyen a múzeum. Másrészt fel kell ismerni, hogy milyen módon lehet kiaknázni a rohamosan fejlődő digitális technológiát, valamint az online elérhető tartalmakban rejlő lehetőségeket.

A fentiek alapján látszik, hogy a múzeumlátogatás kiterjesztése a digitális technológiák múzeumi alkalmazásának legösszetettebb és a múzeum jövője, társadalmi beágyazottsága szempontjából legfontosabb terepe. Ez magyarázza, hogy a digitális múzeum szerteágazó területei közül a saját kutatási-fejlesztési tevékenységemben éppen erre fókuszáltam, és a nemzetközi gyakorlatban is e terület áll jelenleg az érdeklődés középpontjában. A továbbiakban a digitális múzeum említésekor erre a területre szorítkozom. Természetesen van átfedés a különböző feladatokra alkalmazott technológiai megoldások között, és maguk a feladatok is összefüggenek – például egy digitalizált és jól nyilvántartott állomány alapja lehet a kiállítótérben vagy online használható, az interpretációt, oktatást célzó egyes alkalmazásoknak.

A digitális technológiák múzeumi alkalmazásának elvi kérdéseiről, így az információs társadalomban felnövő generációk igényeiről és szokásairól, a modern pedagógia releváns téziseiről, a részvételi múzeum mibenlétéről, a digitális alkalmazások látogatói funkcióiról a falakon belül és kívül, valamint a tervezés és megvalósítás fázisairól, az ehhez elengedhetetlen újfajta interdiszciplináris együttműködésről és múzeumi szerepekről kollégáimmal több alkalommal értekeztünk már.²⁷ Egy esettanulmányban pedig egyetlen alkalmazásunkat elemeztük kritikai jelleggel, pilot látogatói viselkedésfelmérést és véleménykutatást is végezve.²⁸ A cikk hátralévő részében a MOME TechLab Digitális Múzeum alkalmazásaiból válogatva mutatom be a támogatható múzeumi funkciók és technológiai megoldások sokféleségét.

A TechLab alapítását több évtizedes hazai akadémiai pályán töltött időszak, majd külföldi informatikus kutatói és egyetemi oktatói tevékenység előzte meg. A hollandiai egyetemi tevékenységem idején már több projekt keretében foglalkoztam művészeti, művészettörténeti területtel. A műszaki irányultságú hollandiai University of Twente-

²⁶ Bényei Judit és Ruttkay Zsófia, „A múzeum megújítása a digitális technológiák korában,” in *Határtalan médiakultúra*, szerk. Antalóczy Tímea és Pörzsi Zsuzsanna (Budapest: Wolters Kluwer Complex Kiadó, 2015), 51–80, http://buvosvolgy.hu/dokumentum/106/antaloczy_hatartalan.pdf; Domenico Sturabotti and Romina Surace, eds., *Museum of the Future: Insights and Reflections from 10 International Museums* (Roma: Mu.SA Project, 2017), <http://www.project-musa.eu/wp-content/uploads/2017/03/MuSA-Museum-of-the-future.pdf>.

²⁷ Bényei és Ruttkay, „A múzeum megújítása,” 51–80. Bényei Judit és Ruttkay Zsófia, „QR kód helyett digitális stratégia,” in *Digitális Múzeum*, Múzeumi iránytű 12, szerk. Ruttkay Zsófia és German Kinga (Budapest–Szentendre: MOKK–MOME, 2017), 57–74.

²⁸ Ruttkay Zsófia, „ColourMirror: Visitors Identifying with Museum Exhibits in a Playful Way,” in *Kultur und Informatik: Mixed Reality*, Hg. Carsten Busch, Christian Kassung und Jürgen Sieck (Glückstadt: Verlag Werner Hülschbuch, 2017), 51–65.

n – a nemzetközi *hibrid* oktatási kísérletekhez csatlakozva – egy interdiszciplináris szakértőcsapattal létrehoztunk CreaTe néven egy olyan alapképzést (BSc), melynek keretében mérnöki, tervezői, sőt művészi érdeklődésű és készségű fiatalokat képeznek arra, hogy a technológiát humán és művészeti területeken is tudják alkalmazni.²⁹ A képzés – a kezdeti kételyeket megcáfolva – a mai napig nagy sikerrel működik, és azóta számos külföldi egyetemen találni hasonló célú képzéseket. Ilyen jellegű kísérleti, interdiszciplináris munkát jelent a 2010-ben létrehozott TechLab is,³⁰ mely lehetőséget nyújt arra, hogy – oktatási, valamint K+F-projektek keretében – a tervezők társadalmi érzékenységüket és kreativitásukat, valamint vizuális alkotói készségeiket a digitális technológiák eszköztárát kihasználva kamatoztassák.

A Digitális Múzeum portfóliómunkái részint egyetemi keretek között, részint megrendelésre készültek. Egy részük a MOME, valamint más egyetemek (kezdetben az ELTE, az utóbbi hat évben a BME) programozó hallgatói számára meghirdetett (a fent említett) *Digitális Múzeum* című kurzus keretében létrejött, prototípus szintjéig megvalósított és több esetben múzeumi környezetben egy-egy rendezvény keretében ki is próbált projekt. A kurzus tematikája is mutatja a szolgáltatások tervezésének és fejlesztésének összetettségét: a hallgatók projektmenedzsmenttel kapcsolatos, tartalomfejlesztői, vizuális tervezői és programozói feladatokat egyaránt kapnak. A hallgatóknak múzeumi környezetben egy-egy partnermúzeum problémafelvetéseire kell digitális technológiát használva megoldást javasolniuk, valamint együtt kell működniük a múzeumi szakemberekkel.³¹ A bemutatásra kerülő projektek másik részét múzeumi felkérésre a TechLab köré szerveződött TechLab Fellows alkotói kollektíva tagjai készítették, egy-egy időszakos vagy állandó kiállításhoz több interaktív alkalmazást fejlesztve. Néhány megrendelés pedig hallgatói prototípus nyomán született meg, illetve van folyamatban.

2. Esettanulmányok: a múzeumlátogatás kiterjesztése

A továbbiakban több Digitális Múzeum-projektet mutatok be röviden. A válogatás nem a technológiákat, hanem az általuk támogatható funkciókat, az interpretációs lehetőségek sokféleségét reprezentálja. Az egyes projektek alkotói, valamint további részletek (például videók) a hivatkozott online felületen találhatók meg. A hazai gyakorlatról egy frissen megjelent kiadvány ad számot bővebben.³²

²⁹ Ruttkay Zsófia and Ton Mouthan, „CreaTe: A New Programme to Attract Engineers as Design Artists,” *Proceedings of 4th IEEE International Symposium on Electronic Design, Test and Applications* (Hong Kong: DELTA 2008, 2008), 23–25.

³⁰ Moholy-Nagy Művészeti Egyetem TechLab, hozzáférés: 2018.05.22, <http://techlab.mome.hu/>.

³¹ Ez a fajta mérnöki-informatikus, humán és művészi készségeket igénylő, valós feladat köré épülő, a hagyományos egyetemi kurzusok keretein túllépő, a STEAM mozaikszóval fémjelzett képzés egyre nagyobb figyelmet kap külföldi egyetemeken is. A mi kurzusunk több nemzetközi szakmai és oktatási fórumon aratott elismerést.

³² German Kinga és Ruttkay Zsófia, szerk., *Digitális Múzeum*, Múzeumi iránytű 12 (Budapest–Szentendre: MOKK–MOME, 2017).

2.1. Érzelmek, meglepetés és rácsodálkozás

A múzeumi interpretációs eszköztárban ma már az érzelmekre irányuló hatás is nagyobb szerepet kap. Az érzelmi azonosulás, a játékos fizikai tevékenység öröme, a meglepetés bevonja a látogatót, felkelti az érdeklődését, segíti az azonosulást.

2.1.1. Gondolatfújás A PIM *Weöres 100* kiállítására azt a felkérést kaptuk, hogy új-fajta módon hozzuk közelebb elsősorban a fiatal látogatókat Weöres Sándor kötetéhez. A tizenegy interaktív alkalmazás³³ legemblematikusabb darabja a „Gondolatfújás.” A látogató egy álló plexilapba vésett pitypanggal találja szemközt magát. Ha bármiféle instrukció hiányában a gyerekkori emlékeket felidézve és a múzeumi viselkedés szokásos protokollján felülemelkedve a látogató megfújja a pitypangot, akkor annak ejtőernyős terméseiként betűk repülnek egy fal felé, és ott – a fújás intenzitása szerint – rövidebb vagy hosszabb verssorokká rendeződnek. Az egyszerre megmosolyogtató és elbűvölő élmény több puszta játéknál: egyrészt hangulatában a felidézett versek emelkedett, filozofikus gondolataihoz illeszkedik; másrészt megfigyeléseink szerint a látogató – fiatal és idősebb egyaránt – el is olvassa az általa elővarázsolt sorokat, amelyre hagyományos eszközökkel (tipikusan: falra kiírt versekkel) sokkal kisebb az esély.



1. ábra. Gondolatfújás

2.1.2. Gyümölcsversek Ugyancsak a PIM *Weöres 100* kiállításán szerepelt és a felnőttek körében is nagy sikert aratott³⁴ a „Gyümölcsversek.” Színes filcből készült kitömött, nagyméretű gyümölcsök heverték egy asztalon. Ha a látogató kézbe vett egyet, akkor az megszólalt, és a költő egy olyan verse vált hallhatóvá (ahol lehetett, Weöres Sándor saját hangján), melyben az adott gyümölcs szerepel.



2. ábra. Gyümölcsversek

2.2. Interaktív játékos felfedezés

A modern pedagógia gyakorlatát követve³⁵ az információ-közlés a múzeumban is történhet játékos módon, a látogató kíváncsiságára és aktív közreműködésére építve.

2.2.1. Liszt utazásai A Liszt Ferenc-émlékév (Liszt 200, 2011) keretében a Művészetek Palotája számára készült installáció egy zongoraszerű bútor kottaállványába épített

³³ *Weöres 100: A megmozdult szótár*, Moholy-Nagy Művészeti Egyetem TechLab, hozzáférés: 2018.05.22, <http://techlab.mome.hu/weores100>.

³⁴ Papp Sándor Zsigmond, „Gondolatfújás és verstepintás,” *Népszabadság Online*, 2013. júl. 8., http://nol.hu/kultura/20130708-gondolatfujas_es_verstepintas-1398733.

³⁵ Lásd pl. Graham Gibbs, *Learning by Doing: A Guide to Teaching and Learning Methods* (Geography Discipline Network, 2001), <http://gdn.glos.ac.uk/gibbs/>.

érintőképernyőn azt teszi lehetővé, hogy a látogató maga böngéssze a „világ vándora” útját. A repülőjáratokra emlékeztető megjelenítés és a mai Google-térkép a felhasználóknak természetes közelítés funkcióval egyrészt egy szempillantás alatt láthatóvá teszi, hogy milyen aktívan és hova utazott Liszt Ferenc egyes időszakokban, melyek voltak életútjának fontos helyszínei, másrészt lehetőséget ad a különféle nemzetiségű látogatóknak, hogy felfedezzék, a hazájukban, lakóhelyükön is járt-e a mester, adott-e ott koncertet.



3. ábra. Versfoltozó

2.2.3. Versfoltozó Szintén a PIM Weöres 100 kiállításra készült, az előzőekhez hasonlóan a témához illő fizikai megjelenéssel kialakított technológiai alkalmazás az a nagy nyitott könyv, melyben egy-egy hiányos Weöres-vers jelenik meg. A többségében állatokról szóló versek esetében a margókra került, ott táncoló szavakat a látogatónak a tenyerébe véve, óvatosan kell a megfelelő üres helyre beraknia, mintegy befoltoznia a verset. A digitális technológia (Kinect 3Dkamera) újszerű alkalmazásával egy

jól ismert iskolai zsánergyakorlatot (hiányzó szavak pótlását) fordít át az alkalmazás olyan élménnyé, mely – a technológiának köszönhetően az interakció lassúsága révén is – a versre irányítja a látogató figyelmét, azzal már-már intim kapcsolatba hozva őt, miközben a körülállók ugyanúgy aktív részesei lehetnek a történéseknek.



4. ábra. Dürer mozgásban

2.2.4. Dürer mozgásban A gönci Károlyi Gáspár Múzeum *Biblia Kiállítása*³⁶ számára olyan installáció készült, mely képzőművészeti ábrázolásokkal illusztrálva mutatja be Krisztus életútját. Az installáció alkotójának választása Dürer-metszetekre esett a vizuális, stiláris egység megőrzése érdekében. A technológia itt is a lassú szemlélődésre és a szövegek elolvasására készített. A Dürer-jelenetek enyhén megmozduló szereplői és a kísérőszövegek a diavetítés hangulatát és azok kézi, mechanikus vezérlését idézik.

2.2.5. Zenélő képek A Szépművészeti Múzeum a Digitális Múzeum hallgatóitól olyan alkalmazást kért, mely a félreeső helye miatt kevésbé látogatott németalföldi festők képeit kiállító termeket teszi vonzóvá. A képeken gyakran szerepel zenélő társaság, illetve hangszerek, így ez a téma lett a „Zenélő képek” játékos alkalmazás alapja.³⁷ A látogatónak tablettel kell begyűjtenie a különböző hangszereket a festményekről. Ha megtalálta a festményt, akkor a tableten arról szöveges és képi információkhoz jut, megnézheti a régi hangszerek 3D-modelljét,

³⁶ *Biblia Kiállítás*, Károlyi Gáspár Múzeum és Moholy-Nagy Művészeti Egyetem TechLab, hozzáférés: 2018.05.22, <http://techlab.mome.hu/bibliakiallitas>.

³⁷ Az alkalmazás a MOME 2012 őszi Digitális Múzeum kurzusán készült, Sárosi Róbert (MOME), Czibere Ádám (BME) és Könyves-Tóth Zsuzsanna (Liszt Ferenc Zeneművészeti Egyetem) hallgatók munkájaként.



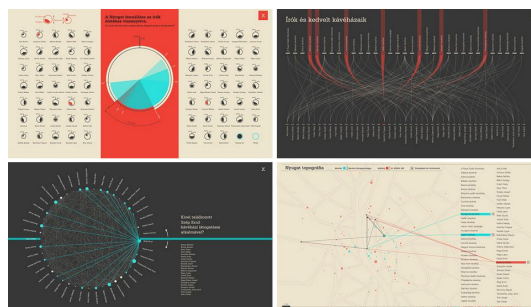
5. ábra. Zenélő képek

sőt, meghallgathatja, hogyan szólhattak a korabeli zenék. A játék lehetőséget ad a hangszerek cseréjére a látogatók között, és azok egyúttal posztolhatók is a közösségi médiában a távol levő ismerősöknek.

2.3. Adatok és összefüggések látható és tapintható módon

Az interaktív adatvizualizáció a múzeumi interpretációnak is hatékony eszköze. Az információ megjelenhet 3D-formában, kinyomtatva és tapinthatóan is.

2.3.1. Nyugat-topográfia Az interaktív képernyőre tervezett, egyetemi diplomamunkaként készült alkalmazás³⁸ a PIM korábbi Nyugat-kiállításához összeállított adatbázist jeleníti meg, teszi az összefüggéseket élvezetessé és könnyen böngészhetővé. A költők, írók és az általuk látogatott kávéházak, azok elhelyezkedése a korabeli Budapesten (képi és szöveges dokumentációval), továbbá hogy ki kivel találkozhatott, illetve egyes személyek és a *Nyugat* folyóirat életének kapcsolódási pontjai jeleníthetők meg interaktív módon. A nagyközönségnek szánt alkalmazás szakemberek számára is új szempontok szerint vizualizálja a folyóirattal kapcsolatban korábban megállapított tényeket.



6. ábra. Nyugat-topográfia

2.3.2. Verstapintó A már korábban említett *Weöres 100* kiállításon a „Verstapintó” nevű installáció³⁹ 3D-nyomtatóval készített hengerek kézbe vételére inspirál. A henger palástjának felületén egy vers ritmikáját és zeneiségét lehetett letapogatni. Tapasztalataink szerint a vizuális és taktilis információ segített felismerni a látogató által választott recitált vers ritmikai mintáit – ezáltal egy egyébként nehéz és sokak számára csak az iskolai tananyagból ismert feladat vált örömteli felfedező tevékenységgé.

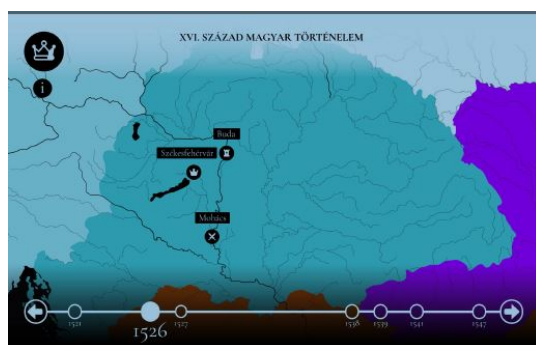
³⁸ Krisztina Szűcs, „Nyugat: Touch Screen Application,” hozzáférés: 2018.05.22, <http://krisztina.szucs.com/my-product/nyugat-2/>.

³⁹ *Weöres 100: A megmozdult szótár*, Moholy-Nagy Művészeti Egyetem TechLab, hozzáférés: 2018.05.22, <http://techlab.mome.hu/weores100>.



7. ábra. Verstapintó

Ebben az esetben is az installáció társas élményt nyújtott nagy méretének és körbeállhatóságának köszönhetően.



8. ábra. Timeline

2.3.4. Timeline Az interaktivitás kétféle módon terjeszti ki a hagyományos papíralapú térképes megjelenítést: egyrészt az interaktivitás révén bevonja a látogatót, másrészt időszímként, a megjelenített információk (például különböző események, helyszínek, népesség és egyéb mutatók) ki-be kapcsolásával információs rétegek, csoportok külön-külön és együtt is láthatóvá válnak. Lehetőség van továbbá a mozgással kapcsolatos jelenségek, folyamatok (például kereskedelem, hadseregek) megjelenítésére.

A Magyar Nemzeti Múzeum történelmi kiállítása számára a *Digitális Múzeum*-kurzus keretében a hallgatók⁴⁰ „Timeline” címmel olyan térképet készítettek, melyen a három részre szakadt ország történelme, szereplői, valamint a mohácsi csata lezajlása válik felfedezhetővé többféle összefüggés mentén, szöveges, képi és animált tartalmakban böngészve.

2.4. Könyvek életre keltve és kibővítve

A könyv mint az ismeretek közlésének hagyományos médiuma – az esztétikai értéket is képviselő fizikai megjelenést és a jól ismert lapozós használatot megtartva – digitális tartalmakkal egészíthető ki a kiterjesztett valóság (*augmented reality*, AR) speciális eseteként.

⁴⁰ Az alkotócsoport tagjai mind egyetemi hallgatók: Dósa Zsófia Lengyel Anna, Rajkó Krisztina, Tick Flóra (MOME); Falvai Olivér, Hornyán Gergely (BME).

2.4.1. Énekes lapozó A Bajor Gizi Színészmúzeum *Kolibri 25* kiállítására hallgatók által készített alkalmazás⁴¹ alapjául a kedvelt és ismert képeskönyv, a *Bors néni* szolgált, mely dalokkal és illusztrációkkal egészült ki. Ha a könyv egy-egy oldalának képét internethez csatlakozott laptop (vagy tablet, okostelefon) kamerája elé helyezzük, megszólal az adott lapon található dal. Egy másik webes felületen pedig az illusztrációkkal lehet „játszani,” a képernyőn végrehajthatjuk azt a cselekvést, ami az egyes dalokban szerepel.⁴²

2.4.2. Bikur Holim A Néprajzi Múzeum *Kő kövön* című időszak kiállításának egyik féltett darabja a „Bikur Holim,” a pápai önszegélyező egyet adománykönyve, mely az adományozóknak csak a nevét tünteti fel egy-egy kis papírzseben. A muzeális könyv természetesen nem fogható kézbe és nem is lapozható, a héber nevek miatt nehezen is lenne értelmezhető a látogató számára. Az interaktív alkalmazás ezeken segít: a könyvet lehet lapozni, az egyes nevekről megtudható, hogy kit takar, valamint családtörténetük dokumentumai is megtekinthetők.



9. ábra. Bikur Holim

2.5. A látogató megszólítása

A hagyományos múzeumi szerep szerint a látogató passzív befogadója a felkínált, szakmai szereplők által összeállított ismereteknek és értelmezésnek. Ezzel szemben a kortárs gyakorlat részvételi szerepet szán a látogatónak.⁴³ Ennek elemi formája az interaktív ismeretszerzés, melyre több példát hoztam a fentiekben. Ezen felül ösztönözhető a látogató önálló véleményformálása is.

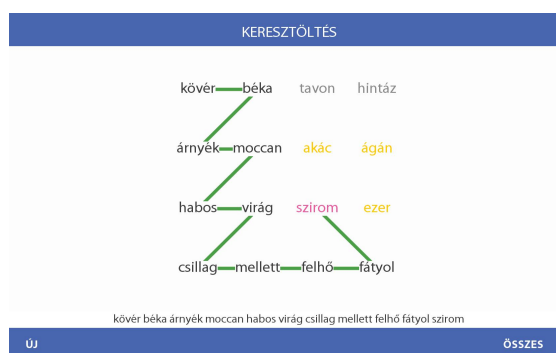
2.5.1. Keresztöltés Az interaktív képernyőn a *Weöres 100* kiállításban a látogató a költő által sugallt többféle olvasási módot maga valósíthatja meg, ha a keresztöltés szabályai szerint hoz létre saját szólancot a (nyomtatásban is) 4×4-es sémában elrendezett szavakból. Azt is eldöntheti, hogy a létrehozott szöveg szerinte versnek minősül-e, azonos rangú-e a költő által javasolt olvasatokkal – ezáltal a vers mibenlétéről

⁴¹ Moholy-Nagy Művészeti Egyetem TechLab, *Énekes lapozó: Kibővített valóság könyv gyerekeknek*, hozzáférés: 2018.05.22, <http://techlab.mome.hu/enekeslapozo>.

⁴² *Bors néni*, hozzáférés: 2018.05.22, <http://techlab.mome.hu/borsneni/>.

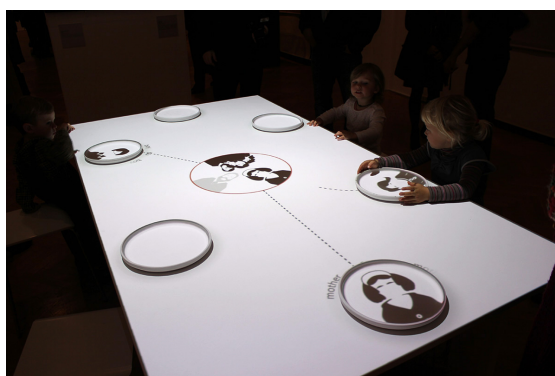
⁴³ Nina Simon, *The Participatory Museum* (Santa Cruz, California: Museum 2.0, 2010).

elmélkedve. A látogatók által előállított versek egy adatbázisban gyűlnek, a kiállítás tartalmát tehát a látogatók is alakítják.



10. ábra. *Keresztöltés*

2.5.2. A szombat köszöntése A *Kő kövön* kiállítás másik installációja a péntek esti családi vacsora szertartását mutatja be. A látogatóknak – a családtagokhoz hasonlóan – helyet kell foglalniuk egy asztal körül, melyre egy animáció vetül a szertartás mozzanatait, az étkeket, valamint az áldások szövegét megjelenítve. Itt a digitális alkalmazás fizikai megjelenése is a tartalomhoz illeszkedik. További új elem, hogy – amiképp a vacsorához is össze kell gyűlnie a családnak – a vetítés is csak akkor indul el, ha legalább hárman leültek az asztal köré. Ez a látszólag megszorító működés éppen a vacsora társas jellegére hívja fel a figyelmet, és egyúttal – a vendégfogadáshoz hasonlóan – résztvevők közötti párbeszédet indukál.

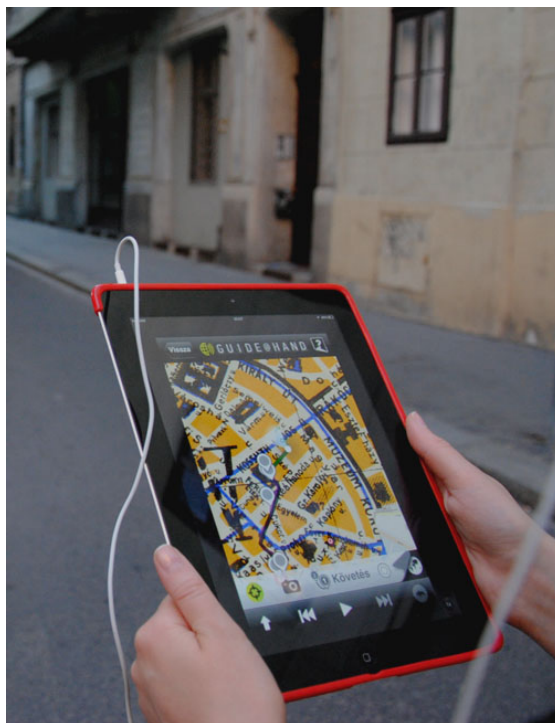


11. ábra. A szombat köszöntése

2.6. A falakon kívül

A múzeum falai között megjelenő emlékek mesterséges környezetben szerepelnek, megfosztva eredeti fizikai környezetüktől vagy kontextusuktól. A kiállítás látogatása-

nak kiterjesztése a falakon kívülre éppen ezt hivatott orvosolni: egyúttal a múzeumba korábban be nem tért közönséget új módon éri el.⁴⁴



12. ábra. *Hajnali háztetők*

2.6.1. Hajnali háztetők A PIM *Ottlik 100* kiállítása számára az egyetemi kurzuson született ötlet továbbvitelével valósult meg a „Hajnali háztetők” című mobil lokatív alkalmazás.⁴⁵ Ottlik Géza a '30-as évek Budapestjén játszódó regényének helyszíneit járhatja be a felhasználó, miközben a regény szövegét hallgatja vagy olvassa. Egy-egy helyszínre érve az irodalmi mű szövege kiegészül korabeli fotókkal, újságcikkkel, a miliőt megelevenítő filmrészletekkel. Lehet idézeteket is kijelölni és továbbküldeni. Az alkalmazáshoz több helyszínen valós élmény is kapcsolódott, például egy Ottlik-idézettel megjelölt pad. Ezáltal ötvöződött a fizikai és a virtuális világ, a mai és a hajdani város. A városi sétát mind történelemtanárok, mind középiskolások vonzóknak találták – és inspirálóknak a regény elolvasására is.

2.6.2. Nézz körül! A „Nézz körül!” weben elérhető alkalmazás⁴⁶ arra példa, hogy az épített környezet miként mutatható be mint a kulturális örökség része. A Salgó-

⁴⁴ Berczi Klára és Mata Juli, „Valós és virtuális séták városi terekben,” in *Digitális Múzeum*, Múzeumi iránytű 12, szerk. Ruttkay Zsófia és German Kinga (Budapest–Szentendre: MOKK–MOME, 2017), 235–247.

⁴⁵ Moholy-Nagy Művészeti Egyetem TechLab, *Hajnali háztetők*, hozzáférés: 2018.05.22, <http://techlab.mome.hu/hajnalihaztetok>.

⁴⁶ *Egy elfeledett budapesti temető: Salgótarjáni úti temető*, Magyar Zsidó Múzeum és Levéltár, Moholy-Nagy Művészeti Egyetem TechLab, hozzáférés: 2018.05.22, <http://zsidotemeto.mome.hu/>.

tarjáni utcában található zsidó temető területén ráadásul igen nehéz a bejárás, és a helyszínen a leromlott állapotú síremlékekről semmilyen információ nem érhető el. A két világháború közötti Budapest szellemi, kulturális, gazdasági és vallási életében jelentős szerepet játszó zsidó családok, személyek síremlékeinek virtuális bejárását teszi lehetővé a „Nézz körül!” alkalmazás. A sírokhoz csatolt szövegek és a korabeli képek mesélik el a családok jelentőségét, valamint az egyes síremlékek építészeti értékeit.

3. Vissza- és előretékinés

A bemutatott alkalmazásokat úgy válogattam össze, hogy reprezentálják a Digitális Múzeum tematikus és technológiai változatosságát, és az alkalmazó partnerintézmények sokféleségét. A tanulmány lezárásaként összefoglalom az általános tanulságokat, és felvázolok néhány jövőbeli feladatot.

3.1. Hol és miért szerepeljen digitális alkalmazás?

Öröndetes az a tendencia, hogy egyre több kulturális örökségi intézmény nyit a digitális technológiák felé. A korábbi, nemzetközi fórumokon is megfogalmazódott kategorikus elutasítást mára az a nézet kezdi felváltani, miszerint korszerű módon élményt adni interaktív digitális eszközök bevetésével lehet és kell is. Ez azonban talán még több veszélyt rejtő álláspont, mint a hagyományokhoz való ragaszkodás. Ugyanis a digitális múzeum eszköztárának mindig a múzeum céljait kell szolgálnia. Ha rácsodálkozik a látogató egy újszerű technológiai megoldásra, netán a múzeumba is azért megy el, hogy azt kipróbálhassa, akkor a technológia játssza a fő szerepet, és a látogató a *kinektezés* vagy a VR-szemüveg élményével tér haza, nem pedig a közvetíteni szánt tartalom, üzenet által megérintve. A saját praxisunkban többször előfordult, hogy igyekeztünk arról meggyőzni a kurátort vagy a múzeumi kommunikációs partnert, hogy az általa elgondolt digitális alkalmazásnak nem lenne kellő létjogosultsága. A múzeum által vizionált megoldás helyett azonban gyakran izgalmasabbat és több célt szolgáló lehetőséget tudunk javasolni. A konceptuális tervezés fázisában fel kell tenni a kérdést, hogy mi olyat ad majd a digitális alkalmazás, amely más módon nem lenne elérhető. A célokat, mint például az érdeklődés felkeltését, az ismeretszerzést és a későbbi felidézést, a célközönség sokféleségét, valamint a használat kontextusát (időigényét, a szükséges készségeket, magányos vagy társas jellegét) mind szem előtt kell tartani. Jelenleg itthon (a sokszor szorító anyagi lehetőségek és időkeretek miatt) ritkán adódik arra mód, hogy több alkalmazást együtt, a kiállítás teljes dramaturgiai ívét, sőt, a múzeum filozófiáját és ahhoz illeszkedő digitális stratégiáját figyelembe véve tervezzünk. Pedig annak is tudatos döntés eredményének kell lennie, hogy a látogatói útvonal mely pontjain (a műtárgyak közelében vagy hangsúlyozottan elkülönítetten) hasonló vagy nagyon eltérő alkalmazásokat szeretnénk-e látni egy kiállításban. Ugyancsak kritikus kérdés, hogy a digitális alkalmazások mellett mennyi időt tölt a látogató a műtárgyakkal, eredeti dokumentumokkal. Véleményem szerint a digitális múzeum eszköztárával le lehet lassítani a múzeumlátogatást, személyre szabott útvonalakat és ismeretszerzési szinteket, illetve különböző megismerési módokat felajánlva rá lehet venni a látogatót az elmélyedésre. Ugyanakkor egy alkalmazás,

az ahhoz összegyűjtött digitális tartalom gyakran több felületen, a falakon kívül is használható; más-más célközönséget, így például a múzeumba el sem jutókat vagy tananyag-kiegészítésként iskolásokat is meg tud szólítani.⁴⁷

3.2. Interdiszciplináris együttműködés

A technológia adta lehetőségek felismerése és kiaknázása nehéz feladat, melyre sem a muzeológus, sem az informatikus, sem a kreatív tervezőművész egyedül nem képes. A kurátor víziója az átadandó üzenetről és a kiállítás koreográfiájáról, a muzeológus (és egyéb szakértő) esetenként sok éves kutatáson alapuló szakmai tudása, a tanulási és interpretációs módszereket inkább ismerő múzeumpedagógus, a gyorsan változó eszközök és technológiák világában járatos informatikus (akik maguk is szakosodnak, például adatbázisok vagy interaktív grafikai felületek programozására), a vonzó interakciót és látványt, a sokszor teljesen újfajta élményt megálmodni tudó vizuális művészek összekapcsolása és együttműködése nélkül nem születhetnek ígéretes múzeumi alkalmazások, melyek sikeressége a kiállításon belüli fizikai beépítésen és térbeli elhelyezkedésen is múlik.

Egy jó alkalmazás megszületéséhez a kreatív tervezőkből, programozókból álló csapat mellé a múzeumi szakma képviselői – kurátor, muzeológus, esetenként egy téma szakértő kutatója, múzeumpedagógus, múzeumi kommunikációs szakember – elengedhetetlenek, ők garantálják a tartalmak hitelességét, és határozzák meg az alkalmazás üzenetét, célját. Igen fontos hogy ki-ki a saját területét képviselje, de legyen nyitott az újszerű interpretációs eszközök iránt, és tudjon a más területről érkezőkkel kommunikálni. A digitális interpretáció nemegyszer külön kutatást (például adatvizualizációhoz elegendő hiteles információ begyűjtését, feldolgozását) igényli. Ugyancsak az együttműködés része, hogy a tartalmakat digitális formában, kellő minőségben és jogtisztán kell rendelkezésre bocsátani. E szűken vett szakmai tevékenységi körökön túl fontos lenne, hogy a tervezés és fejlesztés fázisában is legyenek tesztek beiktatva, amikor a megrendelő és a célzott közönség képviselői visszajelezhetnek. Végül, de nem utolsó sorban: a dizájn, az esztétikum is fontos tényezők. Itt nemcsak egy képernyőtartalom ergonomikus megtervezéséről van szó, hanem arról is, hogy milyen más, az ember számára természetes interakciót tervezhetünk, akár megszokott, a kiállítás környezetébe is illő fizikai tárgyakkal manipulálva, a meglepődésre, játékoságra, érzelmekre is számítva.

3.3. A tudományosság igénye

A digitális múzeum – mint sok interdiszciplináris új terület – még gyerekcipőben jár a tudományosságot tekintve. Ez vonatkozik mind a tudományos elismerés hiányosságára, mind a szakmai megítélés nem meghatározott kritériumaira. A fesztiválok, díjak és a gyakorlatok bemutatására fókuszáló események mellett kevés a tudományos publikációt is eredményező konferencia. A publikációk jellemzően szétszóródnak a hagyományos múzeumi, oktatási, szociológiai és informatikai tudományos folyóiratokban. A szakmaiság és az eredményesség megítélése esetünkben különösen nehéz,

⁴⁷ Vö. Bényei és Ruttkay, „A múzeum megújítása,” 51–80.

hiszen sokszor csak rövid ideig és egy adott helyszínen meg tapasztalható alkalmazásokról van szó, melyek használatának mértékét sok tényező befolyásolja, mint például a kiállítás látogatóinak az összetétele vagy az alkalmazás fizikai helye a kiállítási térben.⁴⁸ Noha egyes technológiák megjelenése a múzeumi térben önmagában sikernek tűnhet, a probléma ennél összetettebb. Az eredményesség mibenlétét a múzeumnak kell többféle dimenzió (például helybeli és online látogatószám, ismeretek és üzenetek közvetítése, a múzeum megítélése) mentén meghatározni. A legfontosabb célok, mint az üzenetek sikeres átadása vagy a látogató attitűdjének változtatása önmagában is nehezen mérhető, akárcsak az, hogy egy kiállítás sikere mennyiben múlt a korszerű interpretációs eszközök bevetésén. Nehéz megbízható összehasonlítási alapot találni, hiszen a digitális technológiák olyan megoldásokat kínálnak, melyeknek nincsen hagyományos párhuzamuk. A használat jól mérhető technikai paraméterei vagy a látogatószám nem adnak képet arról, hogy a használó valóban tanult-e a játékos interakciók során, változott-e az attitűdje, megérintette-e a kiállítás üzenete.

Ilyen kérdések megválaszolása sok empirikus vizsgálatot igényel, melynek metodológiája még hiányzik. A közeljövő egyik sürgető feladata, hogy kialakítsuk a hosszabb távú, interdiszciplináris kutatások számára szükséges intézményes kereteket (például követéses K+F-projektek, PhD-munkák). Ennek számos jelét látni már interdiszciplináris nemzeti és EU-projektek formájában. De maguk a kulturális örökségi intézmények is hoznak létre kutatást célzó formációkat, egyetemekkel együttműködve vagy saját intézményükön belül. A felsőoktatási intézmények új múzeumi, illetve kulturális turisztikai képzéseiben is meg kell azonban jelennie a digitális eszköztár használatának, hiszen a technológia, a társadalmi kommunikációs jelenségek is gyorsan változnak.

Digital Museum – in the Light of the Projects of MOME TechLab

The paper discusses the use of digital technologies in museums. It introduces various connotations of the concept of the digital museum. A special focus will be placed on the potentials of enhancing the museum visits in space and time and engaging the visitor as an active participant. Fifteen projects of the MOME TechLab are introduced, which have been created in an educational setting as student projects or as commissioned works by the TechLab Fellows collective. All of them have been designed for and in cooperation with cultural heritage institutions, representing a broad range of topics from literature to music and fine arts. They are clustered according to mechanisms and goals, such as surprise and emotional involvement, playful exploration, interactive data visualization, augmenting objects with additional information, assigning active role to visitors to voice opinion, and offering real and virtual tours beyond the walls of the museum. Most of the projects are demonstrated at <http://techlab.mome.hu/en>. The paper stresses the necessity of interdis-

⁴⁸ Alicja Knast, „Multimodal Exhibitions: How to Enhance Learning and Design Quality?”, konferencia absztrakt, NODEM 2014 Poland, Varsava. <http://repo.nodem.org/?objectId=335>.

plinary cooperation and the need to scholarly assess the merits of digital installations in museums.

Keywords:

digital museum, virtual exhibition, MOME TechLab

Dragon Zoltán

Szegedi Tudományegyetem, Amerikanisztika Tanszék,

Digitális Kultúra és Elméletek Kutatócsoport

dragon@ieas-szeged.hu

Sebestény Csilla

Szegedi Tudományegyetem, Digitális Kultúra és Elméletek Kutatócsoport

csillasebesteny@gmail.com

#BREW: influencer-kísérlet az Instagram újhullámos kávéközösségében

A #BREW egy közösségi média kísérletnek indult 2015 őszén, melynek célja az volt, hogy valós, konkrét példán, empirikus bevonódáson keresztül tesztelje egyrészt az Instagram képmegosztó közösségi oldal működési mechanizmusát, másrészt feltárja, miként lehet ezt a mechanizmust arra használni, hogy adott időn belül *influencer*ként tudjon megjelenni bárki. Az újhullámos kávé közösség jól körülhatárolható csoportja alkalmat adott arra, hogy a vizuális esztétika és a csoportidentitás összefüggéseit vizsgáljuk, miközben út nyílt egy alternatív marketinges keret felállítására is, amely a közösségi platform gazdasági oldalára enged rátekintést a működési logika és az algoritmus mechanizmusainak feltárása révén. A kutatás így a digitális bölcsészet és a digitális etnográfia, valamint az online marketing eddig egymással nem dialogizáló területeit köti össze úgy, hogy érvényes megállapításokat tesz a jelen pillanatban a Z- és Y-generáció számára talán legfontosabb közösségi platform működéséről.

Kulcsszavak:

arculattervezés, marketing, digitális etnográfia, közösségi média



1. Bevezető

Kutatócsoportunk 2015-ben indította el az Instagram képmegosztó közösségi platformjának vizsgálatát, melynek folyamán egy meghatározott felhasználói csoport viselkedését, identitásformálót, identitást kifejező képi megnyilvánulásait figyelte meg empirikus, bevonódás útján végrehajtott kutatás segítségével. Ebben az értelemben a célkitűzés a digitális etnográfia, a digitális bölcsészet és az online marketing interdiszciplináris határmezsgyéjére pozicionálta magát, ami egy eddig teljességgel feltáratlan területnek számít mind a hazai, mind a nemzetközi szakirodalomban, szinte bármely területet vesszük is figyelembe. A kijelölt területek közül talán a marketing aspektusa

tűnhet kakukktojásnak, ám az empirikus vizsgálat egyértelműen bizonyította, hogy a jelenkori, gazdasági logika alapján működtetett, ezt előtérbe helyező platformlogika nélkülözhetetlenné teszi ezen aspektus komolyabb feltárását.¹ Úgy véljük továbbá, az online marketing területének bevonása olyan új utat nyithat a közösségi platformok átfogóbb kutatása felé, mely ahelyett, hogy kizárná a finansziális, politikai, gazdasági keretek taglalását, egyenesen kiindulási alapként kezeli azt – lévén hogy alapvetően ilyen jellegű és irányultságú profitorientált cégek állnak a működtetés háttérében. Kutatásunk ebben az értelemben is formabontó, hiszen olyan diskurzusokat igyekszik egy cél érdekében párbeszédre bírni, amelyek meglehetősen távol állnak egymástól.

2015. november 23-án regisztráltuk és indítottuk el a #BREW névvel ellátott virtuális identitás Instagram-fiókját azzal a céllal, hogy egy viszonylag kicsi, zárt, szakmailag jól körülhatárolható közösségbe kerüljünk be úgy, hogy mindeközben a szakma egyik aspektusában sem vagyunk otthon, alapvetően semmilyen kötődésünk nincs az adott területhez, így kizárólag a platform adta közösségi médiás lehetőségeket használjuk fel. Ezzel a projekttel azt kívántuk vizsgálni, vajon a jelenkori közösségi média felületei, egészen pontosan az Y- és Z-generációk számára kiemelkedően fontos Instagram milyen identitásszervező, illetve -kifejező eszköztárat nyújt a felhasználó számára, valamint hogy ezen eszközök segítségével miként integrálható az egyén egy meghatározott közösségi struktúrába. Ennek során vizsgáltuk a képi-esztétikai kifejeződés alakzatait és módzatait, valamint ezzel szoros összefüggésben az algoritmus működésének szabályszerűségeit. Ennek kapcsán igyekeztünk olyan megállapításokat tenni, amelyek kijelölik az egyéni érvényesülés lehetőségeit.

A projekt neve az újhullámos kávékészítési kultúra manuális, filteres alapú kávéitalnak elkészítési módjára, a hashtag pedig természetesen az Instagram (és más közösségi portálok) felületén működő hipertextuális címke, illetve linkstruktúra közkeletű elnevezésére utal. A szegmens kiválasztását egyrészt az indokolta, hogy a megcélzott csoport egy jól elkülöníthető, jól definiálható közösség legyen mind összetételét, mind pedig működését és megjelenési módját tekintve. Másrészt annak érdekében, hogy a kísérlet sikeresen bizonyítsa azt a tételt, miszerint komoly szakmai és marketinges csapat, illetve apparátus nélkül, mindenféle előképzettség hiányában is lehetőség van egy közösségi hálózaton jelentős hálózati kapcsolatrendszert kialakítani, a kísérletet lebonyolítóknak ugyanezen kondíciókkal kellett rendelkezniük.

Az Instagram-fiók elnevezését (@hashtagbrew) a közösségi platform keresőlogikája indukálta: a keresőalgoritmus ugyanis ezen az interfészen egyenrangú szemantikai struktúraként értelmezi a hashtagekből létrejövő hipertext alapú hálózatot, valamint a fiókok nevéből kialakuló felhasználói adatbázist. Ennek eredményeképpen ki lehet aknázni a két linkstruktúra összekapcsolását, amennyiben kellően specifikusan határozzuk meg mind a profilnevet, mind a vezető hashtag használatát, mintegy identifikálva a projektet.

A projekt végrehajtása során mindvégig figyelembe vettük és betartottuk az Instagram platformjának használati feltételeit, ami nem pusztán etikai szempontból alapvető

¹ A platformlogika gazdaságelméleti és politikai kifejtését lásd: Nick Srnicek, *Platform Capitalism* (Cambridge: Polity Press, 2017), 43–44. A fogalom esztétikai és formai reflexióját, és kultúrkritikai adaptációját lásd: Dragon Zoltán, „Platform-mozi: Gazdasági, strukturális és formai konvergencia a hollywoodi kommersz játékfilmben,” *Metropolis* 21, 2. sz. (2017): 38–51.

fontosságú, de módszertani oldalról is eredeti célkitűzéseinket erősíti. Kizárólag olyan eszközöket alkalmaztunk a kutatás során, amelyek az Instagram mindenkori felhasználói szerződésében rögzítetteknek, és a törvényi feltételeknek is megfeleltek. Tehát nem használtunk olyan alkalmazásokat, technikákat, megoldásokat, amelyek segítségével gyorsan többszörösére tudtuk volna növelni a követők táborát. Kifejezetten az organikus növekedést kívántuk vizsgálni, amely egy lassabban kibontakozó, ám jóval kiegyensúlyozottabb, elkötelezettebb követői közösségi kapcsolatrendszer eredményez, és amely egy szinten túl további folyamatos növekedést generál. Ezt a marketing stratégiát a kutatás módszertanából kiindulva *slow marketing*ként definiáljuk a *slow*-mozgalom filozófiájának online marketinggel összekötött, és immáron élesben tesztelt variánsaként.

2. Képi esztétika – közösségi logika

Kísérletünkkel nagyjából párhuzamosan egy más jellegű, kvantitatív, az Instagramról nyert képi adatbázisból dolgozó kutatás is futott, melynek vezetője Lev Manovich, a City University of New York Graduate Center professzora volt. Manovich kutatócsoportja *big data* alapon, a kulturális analitika módszereivel vizsgált különböző mintázatokot, és az Instagram felhasználói közösségének szokásai alapján állított fel különböző kategóriákat, valamint tett megállapításokat az Instagram képesztétika variánsaira vonatkozóan.² Kutatásukban a közösségi média felületét a fotográfia történetének következő lépcsőjeként vizsgálták, és számos revelatív kijelentést tettek, amelyek a mi érdeklődésünkre is számot tartanak.

Egyik ilyen jellegű meglátásuk az Instagram alapvető tetszésnyilvánító közösségi eszközére vonatkozik, amely egy szív alakú ikon: ennek aktiválásával tudjuk a poszt szerzőjének tudtára adni, hogy kedveljük a képi tartalmat. Ez Alise Tifentale szerint a kompetitív fotográfia eklatáns példája, amely azonban korántsem csak az Instagram-hoz vagy más, képmegosztó, közösségimédia-jelenséghez köthető. Meglátása szerint:

A kompetitív fotográfia mindig is közösségi, társasági természetű volt. A kiállított vagy a könyvekben megjelent fényképeket más fényképészek nyíltan támogatták. [...] [E]z a fajta fényképészet [...] szorosan követi egy választott képi paradigma tankönyvi előírásait és konvencióit. Vagy, más szavakkal, a kompetitív fotográfia fő ismérve a kedvelhetőség [likeability]. Ahhoz, hogy valaki kedvelhető képeket készítsen, követnie kell a szabályokat.³

A definícióban szereplő, tetszésnyilvánításra vonatkozó szóhasználat kifejezetten arra utal, ami az Instagram kultúrájának motorja: a támogatás, a kedvelés és a kedvelhetőség, szerethetőség attitűdjei és esztétikai kategóriái találkoznak a platform algoritmusának logikájával. A megállapítások arra engednek következtetni, hogy a képi ábrázolás konvenciói nagyban befolyásolják az értékelés attitűdjét, bár Tifentale

² Lev Manovich, *Instagram and Contemporary Image* (manovich.net, 2017), <http://manovich.net/index.php/projects/instagram-and-contemporary-image>.

³ Alise Tifentale, „Defining Competitive Photography,” kutatói blog, 2016. jan. 18., <http://www.alisetifentale.net/research-blog-at/2016/1/17/work-in-progress-defining-competitive-photography>.

azt is megjegyzi másutt, hogy egy-egy tökéletesen tradíciót tagadó, atipikus képi megnyilvánulás is lehet sikeres, elfedve a hagyományos, konvenciókat szem előtt tartó, azoknak megfelelő kompozíciók tömegét.⁴

A kompetitív fotográfia további érdekessége az, hogy a fényképészet történetében, amely jellemzően két részre, az avantgárd és az amatőr vonulat vizsgálatára összpontosít, gyakorlatilag az átfedést jeleníti meg,⁵ vagyis egy olyan határvidék, amelyet az Instagram jelenlegi felhasználói összetételére lefordítva talán a professzionális és az amatőr közötti egyre szűkülő mezsgyének feleltethetnénk meg. Saját projektünk is ezt világítja meg: a kifejezetten amatőr módszerek felől egyre inkább a félprofesszionális felé kellett elmozdulnunk ahhoz, hogy a kijelölt szubkultúra képi kompozíciós és esztétikai tendenciáinak maradéktalanul meg tudjunk felelni.

Manovich kutatócsoportja a rendelkezésre álló minta alapján három fő kategóriát különített el: a hétköznapi, a professzionális és a dizájn fotóké.⁶ Ezeken belül természetesen viszonylag nagy szórás lehetséges a felhasználók társadalmi, kulturális, illetve esztétikai értékei alapján, melynek léteznek helyi és demográfiai aspektusai is. Ennek ellenére megállapítható, hogy a hétköznapi – számunkra legkevésbé releváns – képi stílus Richard Chalfen megfogalmazásában „otthoni módban” működik, vagyis tematikája a család, és elsősorban az ünnepeket, születésnapokat, utazásokat dokumentálja.⁷ Az Instagramon a fentebb említett kategóriák megoszlása a következő: hétköznapi 80%, professzionális 11%, míg a dizájn fotónak minősülő posztok aránya 9%. Előfordulhat, hogy más merítés esetén a pontos százalékok változhatnak, de arányait tekintve megkockáztathatjuk a kijelentést, hogy szignifikáns eltérést nem fogunk tapasztalni.

A #BREW projektje tehát azon túl, hogy szubkultúra tekintetében egy jól definiálható és azonosítható kisebb csoportot célzott meg, vizuális esztétika tekintetében is a legkisebb halmazban pozicionálta magát. Éppen ezért e helyütt nem kívánunk foglalkozni sem a hétköznapi, sem a professzionális fotográfia esztétikai és stratégiai elemzésével, csupán annyit jegyzünk meg, hogy a professzionális és a dizájn kategória között van némi átfedés. Technikai értelemben sem olyan éles a különbség: egyre ritkábban fordul elő, hogy egy kimondottan dizájnprofil mobiltelefonnal készítené a képeit, a többség már elmozdult a rendszerkompaktok, a tükörnélküli digitális gépek felé, amely látványos képminőségbeli javulást eredményezett. Ezt a lépést nekünk is meg kellett tennünk, ezért fél évvel az indulást követően elkezdtünk tudatosan dizájn-fotográfiai elvek mentén posztolni, amelynek köszönhetően folyamatos növekedésnek indult a követőtáborunk. A telefonos posztolással szinte megegyező beállításokat, szituációkat használva arra az eredményre jutottunk, hogy a minőségbeli különbség nagyban befolyásolja a tetszésindexet. Kimondható tehát, hogy ma már egy alapvetően amatőr jellegű közösségi oldalon sem mindegy, milyen képi esztétikával jelennek meg a posztok.

⁴ Alise Tifentale, „Art of the Masses: From Kodak Brownie to Instagram,” *Networking Knowledge: Journal of the MeCCSA Postgraduate Network* 8, 6. sz. (2015), <https://ojs.meccsa.org.uk/index.php/netknow/article/view/399>.

⁵ Tifentale, „Defining Competitive Photography.”

⁶ Manovich, *Instagram and Contemporary Image*, 42.

⁷ Idézi Manovich, *Instagram and Contemporary Image*, 30.

Míg a professzionális fotók esztétikájára jellemző a mélységelesség variálásának maximális kihasználása a téma szerint (tájképeknél nagy, közeli képeknél kis háttérelmosást eredményezve), addig a dizájn fotók kifejezetten egy beállítással dolgoznak, és a legritkább esetben használják a gyújtótávolságban rejlő lehetőségeket. A profik térkihasználásával szemben a dizájn kifejezetten kétdimenziós felületet preferál, az előbbi szimmetrikus szerkesztésmódjával ellentétben az aszimmetriát preferálja. Mindez Manovich meglátása alapján azt jelenti, hogy a dizájn fotó stilizáltabb látványvilága a modern fotográfia és grafikai dizájn alapelveit ötvözi.⁸ Természetesen ezek a megállapítások a *big data* kalkulációk alapján a markáns elkülönítést célozzák. Saját tapasztalataink azt mutatják, hogy a minőséget biztosító félprofesszionális-professzionális, tükörreflexes (DSLR), illetve tükör nélküli (MILC) fényképezőgépek adta lehetőségeket (háttérelmosás, gyújtótáv variációk, mélységelesség, objektívek átfogóképességei stb.) mindkét kategória használja.

Az egyik legeklebtársabb példája ennek a változásnak a VSCO (*Visual Supply Co.*) alkalmazása köré szerveződő közösség fő esztétikai vonala. A VSCO-t 2011-ben alapították Oaklandben, eredetileg az *Adobe Lightroom* fotós utómunkaszoftverhez készítették előre gyártott filmemulációs beállításcsomagokat, majd később mobil alkalmazást fejlesztettek saját szűrőkollekcióval. Idén ismét elővették eredeti portfóliójukat, és előfizetéses rendszerre állásukkal egy időben a filmemulációs szűrők immáron az alkalmazáson belül is elérhetővé váltak. Ennek jelentősége a digitális fotográfia nosztalgikus irányváltásában van: egyre többen térnének vissza az analóg korszak képi esztétikájához, a filmre fényképezés stílusához és színkezelési módszereihez. A VSCO mind az alkalmazás, mind pedig saját galériájának tekintetében előnyben részesíti a digitális minimalizmus⁹ és az analóg színkezelés esztétikáját, melynek megvalósítását elsősorban a dizájn fotó kompozíciós logikájában látják.

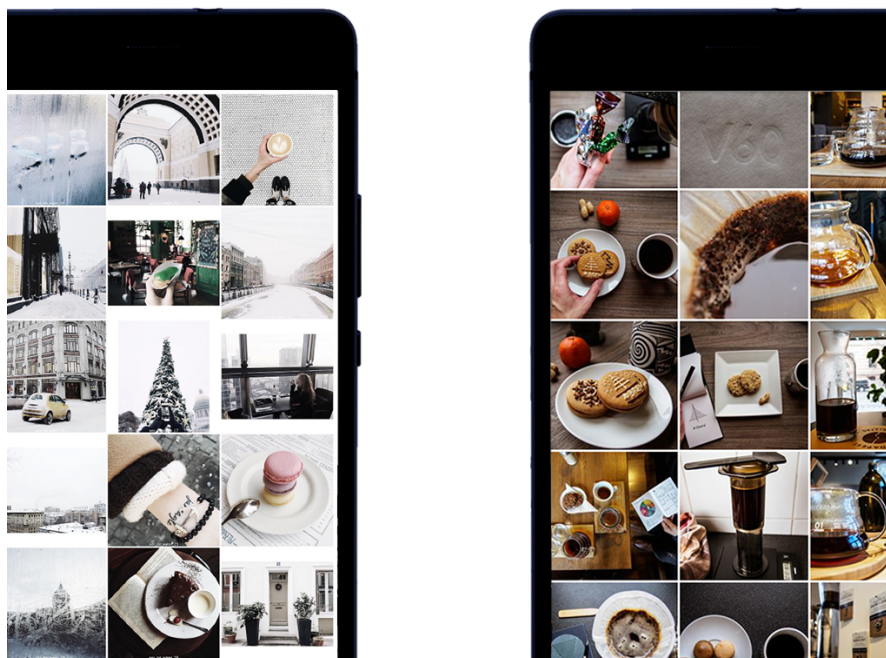
Az Instagram dizájn fotós közösségi terének másik inspirációs forrásaként Manovich a *Kinfolk* magazin köré szerveződő művészeti közeget nevezi meg – joggal. Érdekes módon a *Kinfolk*, csakúgy mint a VSCO, egy évvel az Instagram megjelenését követően indult útjára: 2011 óta vezető életmódmagazin nyomtatásban és online platformon, egyik legjelentősebb tevékenységük a negyedévente publikált tematikus magazin, amely mellett könyveket, eseményeket és kiállításokat szerveznek. A letisztult, skandináv minimalista vonalat követő szerkesztők jellemzően az ezt képileg megfogalmazó pasztell, fakó színvilágot, a szinte végletekig egyszerűsített formai megoldásokat részesítik előnyben. A VSCO hatása mellett e már-már státuszsimbólumként tisztelt magazin vizuális esztétikája köszön vissza az Instagram-felhasználók körében, különösen a dizájn fotók és életképek szín- és formai harmóniájában.

Míg nagy vonalakban valóban egyértelműen megfigyelhető ez a korai esztétikai hatás a VSCO-, illetve a *Kinfolk*-közösségek köreiben, saját tapasztalataink alapján nem osztjuk Manovich összegzését e tekintetben. Már csak azért sem, mert amint az a szabadon hozzáférhető képeken, kollekciókon, mozgóképes anyagokon, illetve a képmódosító szűrők fejlődésén egyértelműen látható, mind a VSCO, mind a *Kinfolk*

⁸ Manovich, *Instagram and Contemporary Image*, 67.

⁹ Az okostelefonok és tabletek képernyőjére, valamint az applikációk használatára kialakult esztétikai válaszreakció, melynek főbb jellemzői a bátrabb vonalak, nagyobb betűk, képernyőt kitöltő képmegjelenítések.

képi világában folyamatos változás jelenik meg. Ennek kapcsán fenntartásokkal kezeljük Manovich kutatásának meglehetősen monolitikus definícióját, amely az általa „instagramizmusnak” nevezett trendet azonosítja.¹⁰



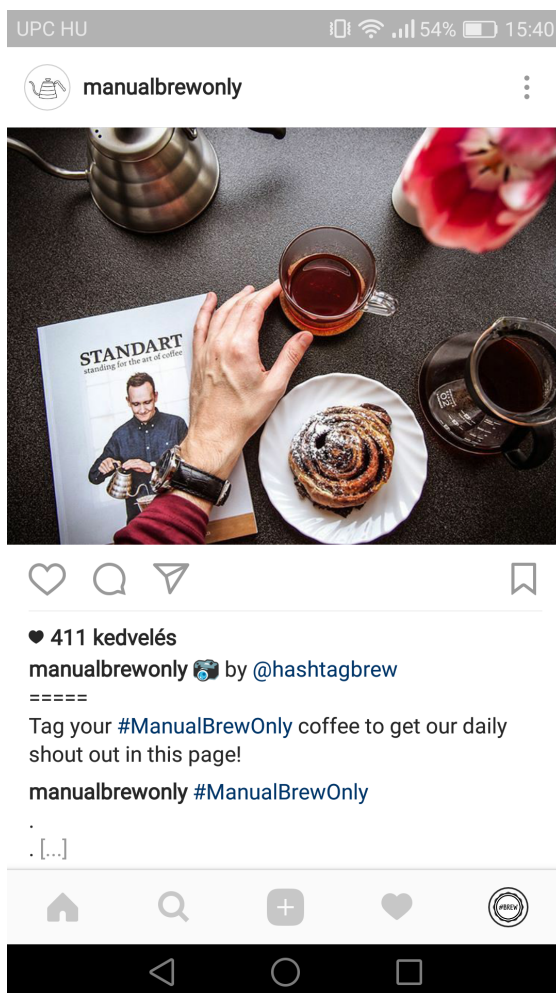
1. ábra. Balra a *Kinfolk* stílusának megjelenési módja, jobbra a #BREW esztétikája

A *Kinfolk* magazin verziója az újhullámos kávé szcénában is fontos és erős jelenléttel bír, ám ez leginkább társadalmi státusz, életvitel, életstílus, a városi hiphop esztétika, illetve annak jelölőjeként jelenik meg (lásd a *@kinfolkandcoffee* profilját, amely explicit módon szervezi *grid*-jét a téma köré). Esztétikai, képi kompozíciós, színharmónia vagy utómunka eljárás tekintetében azonban ez a hatás az esetek túlnyomó részében eltűnik. Ennek igen könnyen belátható oka van, hiszen a kávé színe, a kávézás hangulata, a *pour over*, manuális kávékészítés esztétikája sok tekintetben éppen ellentétes a *Kinfolk* fakított, kissé mosott, pasztelles, minimalista dizájn-jával. Érdekes módon két irányvonal kezd teret hódítani: az egyik a *Kinfolk* mentén pozicionált magazinkultúra kávé irányzata (pl. *Drift*), valamint az esztétikai értelemben újat mutató, de a *Kinfolk* tematikai világából merítő újhullámos, szakmai és rajongói témákat egyszerre felvonultató vonal (pl. *Standart*). Ezek a magazinok egyre inkább átveszik a *Kinfolk* vizuális dizájnmetodológiáját, ám színvilágukban, képi megjelenésüket tekintve eltérnek attól, mert esztétikai értelemben a kávézás kultúrájából építkeznek.

Bár Manovich maga is valamiféle közösséget igyekszik definiálni, amely esztétikai alapon szerveződik, a #BREW-kísérlet ebben a tekintetben, az újhullámos kávé közösségben pontosan az általa leírtak inverzét bizonyította. Manovich Dick Hebdige *Subculture: The Meaning of Style* (1979) című kötete alapján határozza meg a stílus mentén kialakuló szubkultúra képét, ám Hebdige meglátásával ellentétben a szubkultúra

¹⁰ Manovich, *Instagram and Contemporary Image*, 71–95.

az Instagram esetében korántsem ellenkultúraként fogalmazódik meg.¹¹ Az Instagram közösségének vezéralakjai a mainstream mentén, azzal együtt alakítják stílusukat, kísérleteznek kifejezési formáikkal, amelyekbe sok esetben a platformot működtető hirdetési kultúra is jelentős mértékben beleszól.¹²



2. ábra. Rituális esztétika – képernyőmentés a #BREW Instagram oldaláról

Manovich meglátása véleményünk szerint két problémát is felvet. Egyrészt a korábban felállított tipológián belül nem enged meg rétegződést és szóródást, vagyis

¹¹ Manovich, *Instagram and Contemporary Image*, 137.

¹² A szubkultúra kortárs értelmezési lehetőségeit részletes áttekintésben vázolja: Kacsuk Zoltán, „Szubkultúrák, poszt-szubkultúrák és neo-törzsek,” *Replika* 53 (2005): 91–110. A cikk Hebdige szemiotikát is beemelő megközelítését külön kiemeli (98). Manovich Hebdige-kritikája éppen a poliszémia lehetőségeit szem előtt tartva már a poszt-szubkultúra vizsgálatának tárgykörébe tartozik (109), ám mindezen meglátásokat e helyütt csupán említés szintjén kívánjuk beemelni a diszkusszióba, lévén hogy az említett kutatás elsősorban a zenei szcéna alakulására reflektál, és nem tudja figyelembe venni a fotográfiai kifejezőmódok ettől eltérő változásait. Mindenesetre a stíluson keresztül definiált szubkulturális identitás fogalma természetesen termékeny fogalomként működtethető.

még csak jelzés szinten sem árnyalja az egyes kategóriák összetettségét és szerteágazódását. Egy dinamikusan változó közösségi platform esetében úgy látjuk, éppen a folyamatos változás, a kategóriák és műfajok átjárhatósága és egymásra gyakorolt hatása az, ami egybecseng az interfész logikájával és a felhasználói aktivitással. Ehhez szorosan kapcsolódik a másik probléma: a #BREW elmúlt két és fél éve arra enged következtetni, hogy ellentétben azzal, amit Manovich állít, miszerint a stílus határozza meg a közösséget, az újhullámos kávék közegében ez pont fordítva tapasztalható. A vizuális egység, a stílus kodifikációja nem önmagáért való identitásformáló tényező, nem egy korábbi stiláris-esztétikai előzmény lekövetése csupán, hanem magában az újhullámos, filteres kávékultúra rituális mivoltában gyökerezik.

Mindez azt jelenti, hogy az általunk vizsgált szegmensben legtöbbször a kávékészítés folyamata, az eszközök által sugallt esztétika és kompozíció a vezérfonal a vizuális megjelenítésben és tematikában. Véleményünk szerint ez indokolja, miért nem tagozódik be ez a szubkultúra a Manovich által megfogalmazott *Kinfolk*-esztétika képi rendszerébe. Ezen túlmenően az újhullámos kávé esetében a márkák és a termékek önmagukban kávékészítési módokat és technikákat is jelölnek, így nem választható szét a termék és a kultúra sem, ami tovább erősíti a szubkulturális jelölőrendszer működését, hiszen a kávé mint ital, a fogyasztás, az előállítás, és a brandek is szerves részét képezik annak az egységnek, amely az Instagram ezen közösségének képi esztétikáját formálja.

Könnyen magyarázható ez a gyakorlaton alapuló esztétikai formarendszer az antropológia felől, amennyiben Victor Turner rítuselméletét vesszük alapul. Turner szerint a népi kultúrákban megjelenő szertartások, rítusok a társadalmi kohézió alapját képezik: ezek a szimbolikus gyakorlatok válságokon át is képesek egy adott társadalom fennmaradását biztosítani.¹³ Turner elméletével egybecsengően Mary Douglas számára a szimbólumok csoportidentitásokat és értékeket fejeztek ki.¹⁴ A két antropológus meglátását az újhullámos kávék közösség Instagramon való működésére lefordítva tehát azt mondhatjuk, hogy a rituálék – vagyis a kávé készítése és fogyasztása, az újhullámos kávékultúra tárgyi szimbolikája és praxisa – olyan szimbolikus gyakorlatoknak feleltethetők meg, amelyek egyértelműen csoportidentitásokat és értékeket fejeznek ki. Következésképp a #BREW-kísérlet egyik nagy tanulsága az, hogy a kulturális analitika meglátásával szemben nem az esztétika teremt csoportidentitást, hanem a csoport identitásának alapját képező rítusok alakítják a képi kifejezőmódot, a vizuális kohéziót.

A vizuális vizsgálódás mellett egy Instagramon futó projekt sikeréhez szükség van a platformot szabályozó, azt működtető logika feltárására is. Az Instagram eredetileg egy közösségi képmegosztó felületként működött, ahol a követett profilok feltöltéseit időrendileg meghatározott sorba rendezve lehetett megtekinteni, és a saját magunk által feltöltött képek is pontosan ilyen időrendi archívumot, ha úgy tetszik, galériát képeztek. Ez azonban 2016 tavaszán megváltozott, amikor a Facebook által kidolgozott és alkalmazott algoritmikus prioritás alapú megjelenítési rendszer megkezdte működését. Ennek lényege az, hogy a rendszer nem feltétlen időrendben mutatja

¹³ Victor W. Turner, *The Ritual Process: Structure and Anti-Structure* (Chicago: Aldine, 1969).

¹⁴ Mary Douglas, *Purity and Danger: An Analysis of Concepts of Pollution and Taboo* (London: Routledge, 1966), <https://doi.org/10.4324/9780203361832>.

az ismerősök képeit, és nem is feltétlen mindenkiét: súlyozva, prioritizálva kerülnek elénk a képek – ha elénk kerülnek egyáltalán. A legújabb algoritmusváltozat ugyanis, amelyet 2018 elején vezettek be, már többlépcsős megjelenítést alkalmaz. Ez azt jelenti, hogy egy kép csupán a közönség nagyjából tíz százalékához jut el, és ha ebben a szelekcióban sikeresnek bizonyul (vagyis teljesíti a Tifentale által definiált kedvelhetőség kritériumát és kellő mennyiségű tetszést és interakciót vált ki adott idő alatt), akkor egy szélesebb követőbázis elé is eljut.

A kutatás tapasztalata az, hogy egyrészt rengeteg változóval dolgozik az algoritmus, másrészt ezen a platformon is megjelenik az Eli Pariser által definiált szűrőbuborék effektus, ami előbb vagy utóbb, de mindenképpen érezteti a hatását.¹⁵ A buborék felépülését, vagy az abból történő kilépést tudatos stratégiával bizonyos mértékig lehet persze irányítani, ám ez inkább a professzionális felhasználókra jellemző, és finansziális oldala is lehet (hirdetések indítása, kiemelések szponzorálása stb.). Kísérletünk azt bizonyítja, hogy ebben a környezetben alapvetően fontos stratégia a koherens képi tartalom biztosítása mellett a rendszeres publikációs ritmus, valamint a *floodolás* elkerülése (adott időtartamon belüli publikálás visszafogása), mert ezt a gyakorlatot egyértelműen „bünteti” az algoritmus, és hátrányosan befolyásolja a megjelenéseket, eléréseket.

Már a két és fél éves fókuszált kísérlet alapján is azt mondhatjuk, hogy az Instagram platformja néhány év alatt nagy utat tett meg, az egyszerű felhasználók leginkább „pillanatkép esztétikája” uralta képmegosztó közösségi portáljától eljutott a dizájn- és professzionális felhasználók uralta, stratégiai fotográfiára jellemző működéssel felruházott gépezetig. Bizonyos mértékig megmaradt az eredeti koncepció (ti. a felhasználók képeket osztanak meg), ám a hangsúly egyre inkább a monetizációs törekvések felé mozdul el. Először csak direkt hirdetések jelentek meg, majd tulajdonképpen bármelyik fiók tulajdonosa dönthetett úgy, hogy üzleti profillá avanszál, és egyre részletesebb analitikai háttérrel ösztönzik a kiemelések, szponzorált tartalmak gyártására. Mindez magával hozta a nagyobb márkák megjelenését, az *influencer*-struktúra megerősödését és azt a sajátos esztétikai környezetet, melynek egy szubkultúrája révén mi is részesei lettünk.

3. Slow marketing

A #BREW az Instagram adta lehetőségeket igyekezett kihasználni: lassú, tudatos építkezést folytattunk. Kutatásunk nem egy konkrét terméket kívánt eladni vagy pozicionálni, s nem kívántunk monetáris profitra szert tenni. Az átlag felhasználó közösségi integrálódását modelleztük azzal a céllal, hogy véleményvezérré, ún. *influencerré* váljunk, és kiterjesztve az integrációt, közösséget tudjunk kiépíteni, vagyis a hálózati logikát teszteljük a platform működésében. A fenti összefoglalóból kitűnik, hogy az egyéni felhasználó többféleképpen is képes lehet egy adott közösségbe integrálódni (lásd: hashtagek és a vizuális esztétika következetes használata, vagy éppen az

¹⁵ Eli Pariser, *The Filter Bubble. How the New Personalized Web is Changing What We Read and How We Think* (New York: Penguin, 2011).

adott közösség rítusainak adoptációja), azonban világossá vált az is, hogy az énmárka következetes építése is szükséges.¹⁶

A projekt folyamán a márkaépítés alapvető üzenete a következő volt: a kávézásra nemcsak funkcionális szempontból lehet tekinteni, nemcsak arról van szó, hogy reggelente megiszunk valamilyen koffeines italt, hanem fontos maga a rituálé is, az ebből eredő esztétikai élmény, hiszen a használt eszközök önmagukban is pontosan ennek az élménynek a megszületését segítik. Ezek a meghatározások és célok voltak azok, amelyek megteremtették az igényt, hogy egy új megközelítést, ezáltal egy új terminust hozzunk létre, mivel a digitális etnográfában kutatási pozícióként meghatározott immerzív, bevonódás alapú empirika teljesen ismeretlen az online marketing eszköztárában.

A #BREW-kísérlet során így egy saját módszert, a *slow marketing* elvét dolgoztuk ki és alkalmaztuk. A *slow marketing* kifejezésben a *slow* (lassú) szó nem a munka elvégzésének sebességére utal, hanem arra, hogy tudatosabb és fenntarthatóbb építkezésre ösztönözzük az embereket. Maga a kifejezés a *slow* mozgalomhoz kapcsolódik, amely tulajdonképpen egy életfelfogás. Olyan életforma, amelynek keretein belül igyekszünk megtalálni az egyensúlyt, célja, hogy tisztában legyünk azzal, mikor szükséges lassítanunk, vagy éppen gyorsítanunk.¹⁷ Az Instagram közösségének markáns többségét kitevő Y-generáció számára fontos szempont a környezettudatosság, a fenntarthatóság és a minőség – ezzel magyarázható az újhullámos kávé sikere, melynek előállításához hozzátartozik a *fair trade* elve: a hosszú távú fejlődés érdekében jelentősebb összegeket juttatnak vissza a termelőknek, mint általában.

Mindennapos tapasztalat, hogy a fogyasztói igények egyre jobban szegmentálódnak, ami a hagyományos marketingkommunikációt egyre inkább megnehezíti és átalakítja. Az online platformokon létrejövő aktív diskurzusnak köszönhetően azonban ezek az igények teljes mértékben kielégíthetők, így megvalósulhat a teljes personalizáció, amely már önmagában hordozza az egyediség igényét. Az online tér újból helyet adott a kisebb, akár manufaktúrális alapokon működő üzleteknek, és a termékek előállítója és a vevők között egészen baráti diskurzus alakulhat ki, amelytől még inkább úgy érezzük, hogy az az áru személyesen nekünk készült és a webshopoknak köszönhetően az értékesítésnek sincsenek határai. Chris Anderson *hosszú farok* elmélete is erről szól: a tömegtermékekkel ellentétben a szegmentált igényeknek köszönhetően az interneten a kisebb mennyiségben előállított áruk is eladhatók, s mert a készletezés nem az internetes webáruházakat terheli, árubőség keletkezik.¹⁸

A *hosszú farok* olyan szegmentálódást tesz lehetővé, amely begyűri minden lehetséges online platformra, amely valamilyen szinten közösségi elven működik, hiszen ezek a platformok képesek a legnagyobb elérést produkálni. A kutatásban kijelölt csoport tekintetében sincs ez másként, többek között ez is indokolja, hogy az empirikus módszert alternatív online marketinges gyakorlattal egészítettük ki. A #BREW-kísérlet során bebizonyosodott, hogy a technológia, ha megfelelően kezeljük,

¹⁶ John Purkiss és David Royston-Lee, *Énmárka*, ford. J. Füstös Erika és Kónya Rita (Budapest: HVG Könyvek, 2010).

¹⁷ Carl Honoré, *Slow: A lassúság dicsérete*, ford. Dávid Katalin Zsuzsanna (Budapest: Lazi, 2014).

¹⁸ Chris Anderson, *Hosszú farok: A végtelen választék átírja az üzlet szabályait*, ford. Darnyik Judit (Budapest: HVG Könyvek, 2007).

könnyen a céljaink szolgálatába állítható: megteremti az alapot az énmárka felépítéséhez és a közösség tudatos, fenntartható kiépítéséhez a brand köré.

A *slow marketing* módszerének egyik alapvető eleme, hogy a célközönség viselkedési mechanizmusait is vizsgáljuk – esetünkben például azt, mikor kelnek, mikor fogyasztanak kávé, melyek azok a tartalmak, amelyekkel életvitelükből adódóan azonosulni tudnak, milyen nyelvezetet használnak, milyen kulturális elemek határozzák meg a mindennapjaikat stb. –, hogy számukra könnyen értelmezhető és feldolgozható tartalmakat készíthessünk. Többek között ezért gondoljuk úgy, hogy a marketingstratégia kidolgozása és alkalmazása során fontos a humánközpontú szemlélet, a felhasználókra nem lehet csak statisztikai adatokként tekinteni, ahogy ez bevett szokás a marketingstratégiák kialakításakor.

A lassabb és tudatosabb építkezés lehetővé teszi, hogy közösséget építsünk ki a márkánk köré, valamint, hogy olyan dialógust teremtsünk a már meglévő csoportokkal, amelyben saját vizuális-esztétikai identitásunkat az integráció jegyében tudjuk formálni. Ennek folyamán megismerjük azokat a felhasználókat, akik elkötelezett hívei az adott témának, és célcsoportunk véleményvezéreivel is kapcsolatba kerülünk. Így kapcsolódik egymáshoz a tartalom és a marketing: a már korábban említett közösségi logika nagyban meghatározza a stratégiát, amellyel dolgozunk.

Ahogy manapság számos szoftver folyamatosan béta verzióval van jelen a piacon (ez a közösségi médiás platformokra is jellemző), a korábban kidolgozott stratégiánkat nekünk is állandóan változtatnunk kell, reflektálva az újításokra, a célközönség igényeire és elvárásaira. Ez történt például akkor is, amikor az Instagram bevezette az úgynevezett *Instagram story*-funkciót, ami egy huszonnégy óráig elérhető történetmesélési, énközzvetítési réteget teremtett a vertikálisan sorjázó képgalériák fölött, új kihívást és új lehetőségeket adva. Ez radikálisan változtatta meg a platform működési logikáját, hiszen a korábbi, egy-egy képre koncentráló stratégiát felülírta az azonnaliság, a felhasználói bevonódást elmélyítő és szokássá tévő funkció. Egyre inkább tapasztalható tartalomfogyasztási szokás, hogy a kontemplatívabb, elmélyültebb, egyéni képekkel való kapcsolat feloldódik a gyors, mindössze tíz másodperces időtartamú klipek horizontális váltakozásában. Már az egyedi képmegosztások is leginkább a történet funkcióba ágyazva kerülnek a többi felhasználó látókörébe, hiszen ez az első vizuális réteg, amellyel az Instagram-applikációt elindítva találkozunk. Kulturális értelemben véve a lineáris, horizontális történetmesélés még mindig sokkal meghatározóbb, mint a vertikális, újmédia logikán alapuló görgetés.

Kutatásunk azt is bizonyította, hogy az átgondolt, jól felépített, vizuálisan egységes szerkesztett tartalmas és pozitív üzenet elengedhetetlen. Talán naiv meglátásnak tűnik, de a korábbiakban kifejtett kedvelhetőség képi szintű kritériuma a platformon történő viselkedés, etikett szempontjából is igen meghatározó. Bebizonyosodott, hogy az esztétikai élmény mennyire fontos a gasztro- és kávémarketing esetében: a felhasználóknak az étkezés nemcsak létszükséglet, hanem kulturális és vizuális „táplálékként” is szolgál. Ennek során ismertük fel a fentiekben kifejtett meglátást is, miszerint az énmárka felépítésének vizuális vonatkozásait a Manovich körül szerveződött kutatás esztétikai dimenziójával ellentétes pólusból kiindulva szükséges kialakítani. Összességében tehát a *slow marketing* módszerének lényege a bevett, elsősorban időszakos kampányokra épülő marketingstratégiákkal szemben az, hogy egyrészt maximálisan

figyelembe veszi és kiaknázza az újmédiás közösségi platform közösségi potenciálját, másrészt – ennek folyományaként – a fenntarthatóságra törekszik, hosszú távon gondolkodik, nem támogatja a gyors és gyorsan elavuló kampánystratégiákat.



3. ábra. A @manualbrewonly repostja

Az instagramos közösségben nagyon meggyőzőnek számít a követőink száma. Erre technikai oldalon különböző *growth hacking* módszerekkel erősítettünk rá. A fő cél az volt, hogy első körben a megcélzott csoporton belüli releváns követők látóterébe kerüljünk. A *growth hacking* olyan adatközpontú marketingeszköz, amely a tesztelésre épít, így segítve, hogy elérjük a célcsoportunkat. Az Instagram újabb generációs algoritmusának bevezetését követően a bejegyzések nem időrendben jelennek meg a folyamban, ezért fontossá vált, hogy megismerjük, hogyan működik ez az új logika annak érdekében, hogy amennyire felhasználói pozíciónk engedi, befolyásolhassuk azt. Továbbra is fontos a minőség elve, ám egy ilyen közösségi környezetben nem hanyagolható el a mennyiség sem. Ahogyan azt már korábban említettük, a mi esetünkben fontos volt, hogy ne használjunk olyan megoldásokat, amelyeket az Instagram szabályzata tilt, hiszen ez etikátlan, és fals eredményeket hozhat. Olyan alkalmazásokban jelentünk meg a #BREW-val, amelyek *follow-for-follow* elven működnek, vagyis amennyiben be-

követünk valakit, mi is kérhetünk viszonzást. Minden esetben fontos kikötés azonban, hogy a felhasználók maguk dönthessék el, hogy adott esetben mely felhasználónak szavaznak bizalmat, kinek a tartalmaira kíváncsiak. Emellett igyekeztünk a hasonló célközönséggel rendelkező felhasználók aktív követőivel is párbeszédet kialakítani, aminek eredményeként bekerültünk a releváns közösségünk látóterébe is. Így történhetett meg például, hogy a jelenleg 121 ezres létszámú követőtáborral rendelkező, kurátori funkciót betöltő *@manualbrewonly* oldala is átvett tőlünk tartalmat, valamint folyamatosan bekerülnek tartalmaink a közösség által használt hashtagekhez köthető legnépszerűbb képei közé.

Kutatásunk alapján az bizonyosodott be, hogy egy frissen regisztrált felhasználó az Instagram közösségi platformján fenntartható marketing stratégiával, koherens és konzekvens vizuális és esztétikai arculattal nem csupán integrálódni képes egy megcélzott közösségi csoport hálózatába, de a hálózaton belül, a közösségi elveket szem előtt tartva, tudatos építkezéssel *influencer* státuszba léphet. A #BREW-kísérlet tanulsága más vizsgálatokkal szemben az, hogy sem a pusztán marketing oldalú, sem a Manovich kutatása alapján definiált esztétikai koherencia alapján történő pozicionálás nem tud fenntartható eredményt elérni. Ezen meglátás indokolja a különböző területek kombinálását, amely természetesen további kérdéseket és lehetőségeket tár fel a bevont szakmai és tudományos diskurzusok tekintetében.

4. Összegzés

A kutatás eredeti terveinek megfelelően a #BREW-kísérlet első fázisát 2017. március végén lezártuk: ekkorra sikerült ellenőrizhető megállapításokat tennünk a platform működésével, illetve az algoritmus logikájával kapcsolatosan, amelyet felhasználva vázolni tudtuk következtetéseinket mind vizuális esztétika, mind csoport identitás konstrukció, mind pedig marketing szakmai oldalról. A projekt jelenleg stabilan 10000 fölötti követőtábor tudhat magáénak, amely mellé egy tudatosan visszafogott követési stratégia következtében csupán 400 fölötti követés társul.

Eredményeinket több szakmai konferencián és előadás alkalmával is bemutattuk. A kísérlet lezárását követő időszak alatt azonban az Instagram olyan látványos átalakulásnak indult (az *Instagram story* bevezetése és bővítése, a képszerkesztési eszközök és lehetőségek bővítése, a marketinges lehetőségek kiterjesztése, illetve az algoritmus működési logikájának radikális átalakítása), hogy minden egyes összegzésünket folyamatosan adaptálni kellett az éppen aktuális helyzethez. Úgy döntöttünk, folytatjuk a #BREW-projekt működtetését annak érdekében, hogy lehetőségeinkhez mérten mindig rendelkezünk a platformról olyan naprakész információkkal, amelyeket nem elsősorban kutatás, hanem empirikus bevonódás során tudunk tapasztalni és ellenőrizni.

Gyakorlati aspektusai mellett a kutatás további, elsősorban elméleti és kritikai kérdéseket is felvet. Fontosnak tartjuk hangsúlyozni, hogy egy online közösségi platform esetében termékenyebbnek tartjuk azokat a kutatásokat, amelyek figyelembe veszik a platform működésének összetettségét, és külön hangsúlyt fektetnek a gazdasági-financiális motivációk és az ezek mentén kialakuló platformlogika vizsgálatára is. Nem tartjuk elegendőnek azokat a megfontolásokat, amelyek egy alapvetően profitorientált közösségi platform esetében a keretek figyelembe vétele nélkül tesznek megállá-

pításokat szubkulturális vagy csoportidentitás vonatkozásában, hiszen – ahogy azt kutatásunk bizonyítja – sok esetben maguk a körülhatárolható csoportok és azok kulturális jellemzői is a marketinges oldalról megfogható keretek és lehetőségek mentén formálódnak. Ez természetesen súlyos kérdéseket vet fel az olyan, kizárólag teoretikus meglátások ellenőrizhetőségével és relevanciájával kapcsolatban, melyek a kutatásba bevont területek kapcsán általánosabb megállapítások és tendenciák mentén vizsgálják ezen platformok működését. Ennek fényében véljük úgy, hogy bár például Lev Manovich kutatócsoportja fontos általános megállapításokkal gazdagítja az Instagramról való tudományos diskurzust, a kulturális analitika automatizált rendszere nem képes a közösség szerveződésének logikáját megvilágítani, ami a mi nézőpontunkból alapvetően fontos egy online közösségi platform esetében.

Kutatásunk abban a tekintetben is iránymutató, hogy jól elkülöníthető szak- és tudományterületeket kapcsol össze a digitális etnográfiaiban elfogadott bevonódás empirikus eszközének segítségével. Ennek következményeként megnyílik az út az eltérő gondolkodásmódok és meglátások összehangolására, és ezek kooperációjának vizsgálatára. Úgy véljük, a #BREW-kísérlet látványosan demonstrálja ezen diskurzusok kompatibilitásán túl azt is, hogy adott esetben gyakorlati kimenet is azonosítható egy-egy kutatás eredményeire építve.

#BREW: Influencer Experiment in the Third Wave Coffee Community of Instagram

The aim of the #BREW-experiment, started in November 2015, was to test the working logic and mechanism of the social media platform designed for sharing images, Instagram. The researchers opted for an immersive, qualitative approach, focusing on the subculture of third wave coffee enthusiasts (including professionals like roasters, cafés and baristas, along with educated connoisseurs and aesthetically oriented fans). Employing tools developed by digital ethnographers but aiming to explore marketing and visual aesthetics as well, the experiment finetunes and relativizes (at the same time adding a new layer to) Lev Manovich's recently published results that were based on cultural analytic methodology involving big data as a source. The result of our research brings novel approaches both in terms of marketing and digital ethnography, creating an interdisciplinary dialogue with the history of photographic aesthetics and anthropology as well.

Keywords:

branding, marketing, digital ethnography, social media

<KRITIKA>

Maróthy Szilvia

*Eötvös Loránd Tudományegyetem, Régi Magyar Irodalom Tanszék**

mthy.szilvi@gmail.com

Matthew James Driscoll and Elena Pierazzo, eds. *Digital Scholarly Editing: Theories and Practices*. Digital Humanities Series 4. Cambridge: Open Book Publishers, 2016. ISBN 978-1-78374-240-0 (PDF). 294 oldal.
<https://doi.org/10.11647/OBP.0095>

A kiemelkedő nyugat-európai és észak-amerikai digitális filológiai műhelyek kutatói által jegyzett kötet a tudományos elektronikus szövegkiadások legújabb felmerülő problémáival foglalkozik. A monográfia előzménye a hágai Huygens Institute 2012-ben tartott nemzetközi szemináriuma a hollandok mellett olasz, francia, német, norvég, dán, brit, kanadai és amerikai kutatók részvételével zajlott. Mint a szerkesztők megjegyzik, a kötet célja a digitális szövegkiadások diskurzusában fontos, referencia-pontként szolgáló tanulmányok közzététele mind az elmélet, mind a gyakorlat terén.

A kötet elméleti írásainak sorát a szerkesztők közös tanulmánya nyitja meg, mely az utóbbi évtizedek digitális filológiai programjait tekinti át (ennek folytatása Elena Pierazzo tanulmánya is a kötetben). A filológusok tevékenységi körének rövid, adatközpontú meghatározását a kötet egyes tanulmányaira mutatva fejtik ki. A tanulmány szerzői a nem tudományos igényű szövegdigitalizálásról éles véleményt fogalmaznak meg: a nem megfelelően digitalizált szövegek internetes elérése a semminél is rosszabb, hiszen azt a hamis ideát táplálja, hogy a szöveg nem szerkesztési folyamatok végeredménye, az adott kiadás pedig a szöveg egyetlen lehetséges megjelenése (6).¹ Ismerős gondolat ez, a vita könyvtárak és filológiai műhelyek, illetve kalózkodások és könyvtárak között hazánkban is lezajlott, ha nem zajlik jelenleg is (előbbire jó példa a 2004-es *Filológia és digitális barbárság* konferencia anyaga).²

A *digital scholarly edition* (digitális tudományos kiadás, a továbbiakban DSE) fogalmát a kiadástípus első bibliográfusa, Patrick Sahle definiálja, lényegében a 2008 óta fejlesztett katalógusának³ előszavát kibővítve. Meghatározása részint szakmai kritériumokat, részint technikaiakat érint: utóbbi a médiumváltással bekövetkező paradigmaváltás lényegét igyekszik megragadni. A digitálisan létrehozott kiadás e szerint a szöveg struktúráját, jelentését rögzíti, különféle nézeteit (forrásait, verzióit) teszi elérhetővé és vizuális bizonyítékkal (fotóhasonmás) rendelkezik. Az ezzel

* Az Emberi Erőforrások Minisztériuma ÚNKP-17-3 kódszámú Új Nemzeti Kiválóság Programjának támogatásával készült. A kutatás címe: *Elektronikus szövegkiadások Magyarországon*.

¹ Az oldalszám hivatkozások a kötet PDF-verziójára vonatkoznak.

² *Filológia és digitális barbárság: az ELTE BTK Bölcsészettudományi Informatika Önálló Program (BIÖP) konferenciája*, 2004. március 4., hozzáférés: 2018.05.12, <http://magyar-irodalom.elte.hu/biop/barbar/>.

³ Patrick Sahle, *A Catalog of Digital Scholarly Editions*, v 3.0, hozzáférés: 2018.05.12, <http://www.digitale-edition.de/>.

szembeállított digitalizált szövegkiadás pedig az a „hamis” digitális kiadás, mely e tulajdonságokat részint vagy egészében nélkülözi. Sahle a *born digital* kifejezést ugyan nem használja, de világos, hogy a DSE-t ilyennek tekinti. A *reborn digital* fogalmával azonban végképp nem foglalkozik, tehát azzal, miképpen lehet papír alapú kiadásból kritériumainak megfelelő verziót készíteni. Pedig erre is számos példát hozhatnánk: több egyetemi kiadó digitalizálja ily módon korábbi szövegkiadásait.⁴

Amikor azonban a DSE megfogalmazásával Sahle az adott mű összes forrásainak teljes körű feldolgozását és megjelenítését, a szöveg többszörös megjeleníthetőségét hirdeti, valójában nem a digitalizálttal helyezkedik szembe, hanem a kritikai kiadás, a főszöveg-megállapítás eljárásaival. A digitális fordulattal tehát éppoly elfogult, mint volt a nyolcvanas évek szövegkritikája a genetikus szemlélettel kapcsolatban. De miért ne lehetne az elektronikus kiadás akár kritikai, akár forrás-, akár genetikus vagy éppen népszerű kiadás? A DSE mint a lehető kiadások legjobbika, a mindentudó és objektív⁵ naivan hatnak abban a sokkal szabadabb közegben, amit az utóbbi évtizedben nemzetközileg elfogadottá vált Text Encoding Initiative (TEI) szabványa biztosít.

A tudományos elektronikus szövegkiadás mint fogalom magyarázatánál nemcsak az elektronikus (vagy digitális, hálózati) szorul magyarázatra, hanem a kiadás is. Az elektronikus, kivált a mára egyeduralkodóvá vált hálózati publikációk alapvonásai a nyitottság, a tartalom, a technikai kivitelezés és a létrehozói csoport folyamatos változása, ezt számos tanulmány hangsúlyozza. A dinamikusság olyan előnyei mellett, mint a folyamatos fejleszthetőség, a valós idejű kiadás (*edition-in-time*) jelensége, a hipertext lehetőségek vagy interaktivitás számos hátránnyal is jár. Ilyenek többek között a kevés példány miatti sérülékenység, a korábbi verziók felülírásával létrejövő adatvesztés, a nem megfelelően biztosított hivatkozhatóság.

A kötetben Marina Buzzoni egy a kiadástípusra vonatkozó protokoll létrehozását veti fel, s foglalja össze, egy ilyen ajánlásnak miket kellene figyelembe vennie – konkrét javaslatokat azonban nem fogalmaz meg. A TEI kézikönyvében (2007) Marilyn Deegan már kísérletet tett a gyűjtés és megőrzés szempontjainak felsorolására, a kiadások gyűjteményi kezelésére: ez máig az egyik legalaposabb munka a témában.⁶ A TEI egyik munkacsoportja pedig 2015 óta dolgozik a TEI Simple modellen, mely lehetővé teszi, hogy a TEI kódolás ne csak a szöveg jelentésére, hanem annak megjelenítésére vonatkozó információkat is rögzítsen, lényegében TEI elemkészlet használatával.

A tudományos elektronikus szövegkiadások összegyűjtésével, katalógusszerű feldolgozásával foglalkozik Greta Franzini, Melissa Terras és Simon Mahony tanulmánya. Munkacsoportjuk 2016 óta dolgozik a nemzetközi katalóguson, melyet Sahle annotált bibliográfiájával szemben egy jóval több szempontot (49-et) figyelembe vevő és közös-

⁴ Például: Paul Vetch et al., *The Cambridge Edition of the Works of Ben Jonson Online* (Cambridge University Press, 2014), hozzáférés: 2018.05.12, <http://universitypublishingonline.org/cambridge/benjonson/>; Mikszáth Kálmán művei, DigiPhil, hozzáférés: 2018.05.12, <http://www.digiphil.hu/>.

⁵ „Within the digital paradigm, the process is reversed: the editor does not *write* the edited text.” *Digital Scholarly Editing*, 31. (Kiemelés az eredetiben.)

⁶ Marilyn Deegan, „Electronic Textual Editing: Collection and Preservation of an Electronic Edition,” in *Electronic Textual Editing, Preview*, John Unsworth, Katherine O’Brien O’Keeffe and Lou Burnard, eds. (tei-c.org, 2007), http://www.tei-c.org/About/Archive_new/ETE/Preview/mcgovern.xml.

ségileg fejleszthető modell jellemez. Előzetes felmérések szerint 325 tételt kell feldolgozniuk, ebből jelenleg 222 van kész. Alapvető céljuk a katalógussal az elektronikus kiadások követendő gyakorlatának megrajzolása, tehát elsősorban nem a megőrzést, a gyűjteményi feldolgozást tartották szem előtt. Sajnos alapvetően a nyugat-európai és észak-amerikai kiadványok reprezentálják magukat, magyar vonatkozású kiadás egyetlen található, az *Inscriptiones Alborum Amicorum*.⁷

A kötet az utóbbi évek trendjeiből széles merítést ad. Dirk Van Hulle a Beckett archívum vezető kutatója annak elméleti hátterét mutatja be, miként lehet alkalmas egy folyamatosan bővülő szerzői szövegbázis kéziratokkal, nyomtatványokkal és szerzői könyvtárral a szövegkeletkezés vizsgálatára, illetve az intertextusok feltérképezésére. Elemzése a genetikus kiadás határait feszegeti, megfogalmazását kölcsönözve: „a marginália és a paralipomena képes megmutatni a »kiterjesztett elme« működését” (118).

A kiadások sokféleségére, rétegeire mutat rá Krista Stinne Greve Rasmussen is, aki az olvasói, felhasználói és közreműködői szerepeket különíti el és vizsgálja. A közösségi használat nagy hangsúlyt kap a Devonshire Manuscript projektet bemutató sokszerzős tanulmányban is. A projekt célkitűzése az volt, hogy olyan tudományos kiadást hozzanak létre, mely egyúttal közösségi is, bevonva a Wikibooks, a Facebook vagy a Twitter felületeit. A tanulmány erről a munkafolyamatról számol be, és az *open scholar*rá válásról: ne csak a publikáció, a kutató is legyen nyitott – mondják.

A TEI alkalmazásával, tehát a kódolás gyakorlatával a kötetben mindössze két tanulmány foglalkozik részletesen. Az egyik a TEI szabványban a levelezések kódolására nemrégiben bevezetett elemkészlet (*correspDesc*) ismertetése Camille Desenclos tollából, aki további fejlesztési javaslatokat is megfogalmaz. A másik Cynthia Damon alapos dolgozata, mely példákon keresztül mutatja be, az antik szövegek kódolására miért nem használható jelenleg a TEI kritikai apparátusa, s milyen elemekre volna és nem volna szükség. Állítása szerint a TEI kritikai apparátusa a szövegkorpuszra inadekvát, a használatban lévő apparátus összetettebb és kifejezőbb. A digitális átállás azonban egyre sürgetőbb, hiszen antik szövegek kritikai kiadásai még mindig szinte csak nyomtatott formában érhetők el, mely (mást nem említve) az oktatás szempontjából egyértelmű hátrányt jelent.

A kötetben az elektronikus szövegkiadások kritikájával is több tanulmány foglalkozik. Fontos észrevételei vannak Joris van Zundertnek azzal kapcsolatban, mennyire kiaknázatlanul hagyják a hipertext nyújtotta lehetőségeket az elektronikus kiadások. Többségük nem a hipertext modellre épül és egyes elemeit is gyakran alig azonosítják, például ugyanazon URL szolgál azonosításul az adott levél fotójára, átírására és metaadataira. Van Zundert állítása szerint ezen kiadások csupán szó szerinti fordítások könyvszerűből nem könyvszerű médiumba (103).

A könyv talán legizgalmasabb tanulmánya Roberto Rosselli del Turco írása, mely sorra veszi mindazokat a hibákat, melyek egy elektronikus szövegkiadásnak felróhatók. Az elektronikus forrásokhoz a kutatók többsége máig szkeptikusan viszonyul, s ennek fő okozója éppen az a folyamatos változás, melyet a tanulmányok többsége is (inkább pozitívként) kiemelt. Sokszor nem lehet (oldalra, paragrafusra) pontosan

⁷ Latzkovits Miklós, *Inscriptiones Alborum Amicorum*, hozzáférés: 2018.05.12, <http://dx.doi.org/10.14232/iaa>.

hivatkozni, a hozzáférési útvonal gyakran változik, akárcsak a tartalom: nincsenek editiók, egyik verzió felülírja a másikat. Gyakran ad okot aggodalmakra az elektronikus kiadás programozásának, üzemeltetésének magas költsége is. Megoldatlan probléma a kiadások többségénél az eszközre optimalizálás. A tanulmány szerint még a folyamat elején tartunk, még nem vívtuk meg azt a csatát, ami a digitális forrásokat tudományosan elfogadottá teszi.

Összefoglalva, a *Digital Scholarly Editing: Theories and Practices* a TEI által közreadott kézikönyvek és a digitális bölcsészettel, filológiával foglalkozó monográfiák után egy új naprakész munka kíván lenni, mely a tudományos elektronikus szövegkiadások elméletére és gyakorlatára hoz példákat. Bár a kötetnek nem célja átfogó ismeretek átadása, s többen korábbi publikációikat frissítették csak fel, kétségkívül megismertet a szövegkódolás újabb irányjaival, divathullámaival és leküzdendő problémáival.

Maczelka Csaba

Pécsi Tudományegyetem, Angol Nyelvű Irodalmak és Kultúrák Tanszéke

maczelka.csaba@pte.hu

Laura Estill, Diane K. Jakacki and Michael Ullyot, eds. *Early Modern Studies after the Digital Turn. New Technologies in Medieval and Renaissance Studies* 6. Toronto: Iter Press, 2016. ISBN 9780866985574. vii, 378 oldal.

Az Arizona Center for Medieval and Renaissance Studies sorozata 2009-ben indult útjára, és a nyitókötet címe (*New Technologies and Renaissance Studies*) még a digitális technológiák újszerűségét emelte ki. A jelen ismertetés tárgyát képező, 2016-ban megjelent hatodik kötet viszont már címében is hangsúlyozza, hogy a digitális fordulat elvégeztetett, és eljött a számadás ideje. Ennek szellemében az illusztris szerzőket felvonultató kiadvány legfőképpen arra keresi a választ, hogy a koraújkor-kutatás különböző részterületein milyen elméleti-módszertani változásokat hozott a digitális bölcsészet előretörése. Ez az igény ugyanakkor a legtöbb esetben szerencsésen kiegészül a bevezetés szándékával, a szerzők jól láthatóan igyekeznek az érintett területeken kevésbé jártas olvasók számára is áttekintést nyújtani az új technológiák kínálta lehetőségekről, sőt, talán ez a szándék az, ami valamelyest képes egy keretbe foglalni a tematikailag egyébként nagyon is sokszínű, már-már széttartó kötetet. A tanulmányok emellett folyamatosan igyekeznek a befogadót is újabb kérdések megfogalmazására ösztökélni, mert szerintük a technológia jelenleg olyan mértékben jár előttünk, hogy a kimeríthetetlenül sokféle alkalmazhatóság felderítése mindannyiunk felelőssége – a kötet ehhez nyújt megszámlálhatatlanul sokféle megfontolást és segítséget.

A rövidke bevezető az ösztökélés szellemében fogant, így a szokványos összefoglalás helyett inkább egyfajta használati utasítást biztosít az olvasónak, akit a szerkesztők arra biztatnak, hogy az esettanulmányokban tárgyalt konkrét projektek megismerésén túl igyekezzen folyamatosan reflektálni arra is, hogy a különféle megközelítések, eljárások, technológiák, szoftverek és szolgáltatások az ő saját területén hogyan lehetnének alkalmazhatók. Ezt azért is sürgetik, mert szerintük most vagyunk azon a ponton, amikor fel kell mérni és újra kell gondolni a digitális fordulat következményeit. A változás horderejét sokat tanulmányozott kora újkori analógiával illusztrálják: ahogy a teleszkóp alapvetően megváltoztatta a 17. századi ember világgképét, úgy alakítja át ma a kora újkor kutatóinak a koncepcióit a digitalizáció: az újdonság mindkét esetben abban rejlik, hogy korábban még nem érzékelhető dolgok válnak újonnan láthatóvá. A kötet tanulmányai, melyek nagyon sok esetben koncentrálnak valamely újfajta vizualizációs technika köré, nagyon jól illusztrálják ezt az alaptételt.

A tulajdonképpeni nyitószövegben Michael Witmore és Jonathan Hope, a digitális bölcsészet két nagyágyúja nyilvánvalóan azt a feladatot tűzte ki magának, hogy egy nagyobb ívű, összefoglaló tanulmányban közérthetően, párhuzamok és analógiák sorával illusztrálva készítsenek pillanatképet a digitális bölcsészet jelenlegi állapotáról

és kihívásairól. Ez a régi és új nagyon érdekes ötvözésében nyilvánul meg: arra jutnak ugyanis, hogy ahogy az évszázadok során függetlenedett egymástól a nyomtatott könyvek logikai rendszerezése és fizikai tárolása, úgy a digitalizálás is lényegében a többszörös, párhuzamos, hagyományos vagy radikálisabb újrendszerezés lehetőségét teremtette meg (17). Az elméletibb bevezetőt konkrét példák követik, és ezek, nem meglepő módon (Witmore a Folger Shakespeare Library igazgatója) a Shakespeare-korpuszhoz kapcsolódnak. A *king*, a *love* és a *might* szavak gyakoriságának vizsgálatán keresztül a drámák olyan érdekes újrendezéseit figyelhetjük meg, amelyek szinte automatikusan szülnék újszerű kutatási kérdéseket. Miért látszik, a szógyakoriság alapján legalábbis, eltűnni a szerelem a késői drámákban? Vajon tekinthetjük-e általánosabban egy bizonytalanodó, egyre hipotetikusabb (mondhatni: manieristább?) nyelvhasználat tükrének azt, hogy a pályája előre haladtával Shakespeare egyre gyakrabban használja az esetlegességet kifejező *might* segédigét? Az éppen csak felvillantott kérdésektől egy hatalmas ugrással jutunk el a sokkal komplexebb, átfogó irodalomtörténeti relevanciával bíró témákig. A két szerző nem kevesebb, mint hetvenkét úgynevezett nyelvi cselekvéstípust (*linguistic action types*) vizsgál meg a *DocuScope* nevű nyelvi elemzőszoftverrel, és a korpuszt kiterjeszti 554 kora újkori angol drámára. Ebből nem csak azt tudjuk meg, hogy Shakespeare és sokkal kevésbé ismert kortársai között az így mérhetővé tett nyelvhasználat tekintetében nincs számottevő különbség, de felülíródik az a szakirodalmi közhely is, hogy az angol tragédia fő stílári forrása Seneca lenne, miután azt kell látnunk, hogy az ebből a szempontból oly fontosnak tartott korabeli Seneca-fordítások e mérés alapján egyáltalán nem hasonlítanak a kor professzionálisan szabványosnak látszó modelljére. Ebből már felsejlik az általánosabb, nem csak erre a tanulmányra vonatkozó tanulság: a nagyobb számban elérhető, automatizáltan elemezhető szövegeknek köszönhetően olyan új típusú irodalomtörténet körvonalai kezdenek el kibontakozni, amely a kiugró egyéni teljesítmények és a fordulópontok helyett sokkal inkább az irodalomtörténet folytonosságára és a hasonlóságokra figyel.

Mattie Burkert dolgozata több ponton kapcsolódik az előző esszéhez, ám a kutatás tárgya ezúttal a 17. század végének angol drámairodalma. Az addig egyeduralkodó társaság szétválásával 1695-ben két rivális társaság (Drury Lane, Lincoln's Inn) alakult, ami új lehetőséget teremtett a feltörekvő írók számára. A kor drámaszerzőinek harmada női író, akiket a *The Female Wits...* című szatirikus dráma azzal vádol, hogy önteltségükben felelevenítették az 1660-as, 1670-es évek hősi tragédiájának a hagyományát, és azt nőies érzelgősséggel is sikeresen „gazdagították” (35). A szakirodalom nem egységes abban a kérdésben, hogy mennyire volt valóban gender-központú jelenség az 1690-es években újra felfedezett hősdráma, de sokan feltételezik, hogy a forma és az előző generációhoz (Aphra Behnhez, Katherine Philiphez) is tudatosan kapcsolódó női irodalmi hagyomány között fennáll valamilyen összefüggés. Az előző dolgozat módszertanához hasonlóan Burkert is nagy korpuszon (közel háromszáz, 1670–1699 közötti drámán) végez nyelvészeti vizsgálatokat a *DocuScope* segítségével, és így próbálja meg kideríteni, hogy mennyire köthetők a formával kapcsolatos időbeli változások a szerzők neméhez. Az igen részletes technikai ismertetést tartalmazó dolgozat végső konklúziója az, hogy bár a korpuszban a három évtized alatt jól kimutatható nyelvi változások történnek, ezek valójában nem köthetők női szerzőkhöz. Vagyis a

szakirodalom jól látta, hogy a kor női szerzői sajátos stílusban írtak, de nemcsak ők írtak abban, és ha ez a stílus bármennyire is köthető nemekhez, akkor nem a szerzők, hanem inkább a szereplők nemével látszik összefüggésben lenni. Mint Shakespeare és kortársai esetében, itt is azt tapasztaljuk, hogy a kvantitatív nyelvi változások régi szakirodalmi beidegződéseket kérdőjeleznek meg.

Maciej Eder tanulmánya a holt nyelvek digitális tanulmányozásának lehetőségeit mutatja be egy elgondolkodtató korpuszvizsgálati példával. Az általa alkalmazott stilometriai módszer történeti dimenzióira is figyelmeztet: a módszer alkalmazásának legismertebb példájaként kerül elő a humanista Lorenzo Valla neve, aki lényegében a stilometria segítségével bizonyította, hogy az Itália területét állítólagosan a pápának adományozó constantinusi adománylevél nem születhetett a 4. században. A tanulmány azonban kifejezetten arra törekszik, hogy a szokványos statisztikai alapú attribúciós eljárásokon túlmenően mutassa be a digitális technika alkalmazhatóságát, mégpedig úgy, hogy megvizsgálja a kora újkori latin irodalom viszonyulását a korábbi irodalomtörténeti korszakokhoz. Ehhez százötven, mellékletben felsorolt szövegben vizsgálja a leggyakrabban előforduló szavakat, melyek három nyílt hozzáférésű adatbázisból származnak, és ókori, ókeresztény, középkori és kora újkori szerzőket egyaránt tartalmaznak. Kulcsfontosságú az adatok feldolgozásában a vizualizáció – a dolgozat legfontosabb megállapításai lényegében a stilometria segítségével feltárt, korábban nem ismert viszonyok felismeréséből származnak. A vizsgálat ugyanis a legközelebbi szomszéd (*nearest neighbors*) módszerét alkalmazva vázolja fel a különböző korokban alkotó szerzők közötti stílári kapcsolatok hálózatát, és ez a hálózatos megközelítés hol megerősít korábbi feltételezéseket, hol pedig egészen meglepő új eredményekhez vezet. Az egyik átfogó következtetés az, hogy a latin stílus nem lineárisan fejlődik, különösen a kora újkori szövegek esetében figyelhető meg nagyfokú variancia – hol ókori, hol kora vagy késő-középkori szövegekhez állnak közelebb. Ennél az általános elvárásainknak jobban megfelelő eredményt hoz ugyanazon mátrix műfajok szerinti újracsoportosítása, amelyből jól kivehető, hogy a műfajok hajlamosak önálló stilisztikai szigeteket alkotni a nagyobb korpuszon belül. A Ciceró hatását leképező ábra kevésbé meglepő, ám éppen ezáltal a módszer alkalmazhatóságát bizonyító eredményeket hoz, amikor bemutatja, ahogy More és Erasmus anti-ciceroníanus elmélete a gyakorlatban is tükröződik. A tanulmány jól példázza és szorgalmazza azt, hogy a digitális módszerekkel végzett stilometria, és általában a távoli olvasás hogyan segíthet új, a hagyományosabb módszerekkel (stilisztika, szoros olvasás) megválaszolandó kérdések feltérképezésében.

Jesús Tronch tanulmányán keresztül ismét visszatérünk Shakespeare-hez, ezúttal a *Hamlet* válik állatorvosi lóvá egy többváltozatú szöveg többnyelvű hipertext kiadásának esetenként igen mélyen technikai jellegű kérdései kapcsán. A szerző által egy nagyobb projekt (a kora újkori európai színházi hagyománnyal foglalkozó EMOTHE) keretében felvázolt kiadás célja az, hogy túllépve a szövegek pusztán egymás melletti megjelenítésén, közvetlenül segítse is az olvasót az összehasonlításban. Külön figyelmet kap a forrásszövegek és fordítások összehasonlítása, és aki valaha is dolgozott már számítógépes fordítástámogató eszközökkel, az mindjárt sejtheti, hogy a kulcsfogalom a *segmentálás* és az *igazítás* lesz. Tronch projektjében három korabeli angol és hat modern angol változat mellett két spanyol fordítás (ezek közül az egyiknek négy

variánsa), két német, két olasz és két francia fordítás, összességében tehát húsz változat válik összehasonlíthatóvá. A projekt egyik kiemelt célja az, hogy szövegváltozatok esetében felhívja a figyelmet az eltérésekre, vagyis hogy ne egy szövegen belül mutasson be párhuzamos olvasatokat (ami a nyomtatott kiadások bevett gyakorlata), hanem magukat a párhuzamos szövegeket jelenítse meg. Amíg a tanulmány ezeket a megfontolásokat szemlézi és összegzi, addig már szinte gyanúsán kevés az újdonság. A gyanú a dolgozat második felében igazolódik be, ahol különösebb előzetes figyelmeztetés nélkül egyszer csak a szövegfeldolgozó szoftverek mélyvizében találjuk magunkat. Szó esik a *TUSTEP*, a *Juxta*, a *CASE*, a *TEI Comparator* és a *Collatex* nevű kollációs szoftvekről és a *TCA2*, a magyar vonatkozású *Hunalign* és az *InterText* nevű szövegigazító programokról. A programok ilyen széles arzenáljára a vállalkozás hibrid jellege miatt van szükség: bár mind szövegek összehasonlítására és/vagy igazítására szolgál, egyik sem úgy készült, hogy egyszerre lehetővé tenné az egy nyelven belüli változatok, és a különböző nyelvű szövegek közötti összehasonlítást. Mivel a fordításokban gyakran egészen komoly strukturális különbségek is előfordulnak, ezért a beszédjelzések, színpadi utasítások és a beszédek sorszáma sem használható a párhuzamos szöveghelyek összekapcsolásához. A feladatot végül félautomata gépi előszegmentálás és emberi ellenőrzés kombinációjával hajtják végre, és az így összekapcsolt szegmenseket a felhasználó kattintással igazíthatja egymás mellé. Erre viszont a szövegek kódolása során is figyelni kell, így a tanulmány utolsó szakasza az egymástól eltérő helyen található fordításpárok kódolásáról szól, és konkrét részletekkel, majd a függelékben közölt több oldalnyi egybefüggő kóddal segíti a probléma megértését. A tovatűnt tinédzserkor és az utoljára talán 2008-ban megjelent *Programozzunk Turbo Pascal nyelven!* és hasonló kiadványok jutnak az olvasó eszébe – ez a tanulmány, különösen a második fele, nem a nyúl szívűeknek szól, legalább minimális kódolási jártasságot feltételez.

A következő dolgozatban Rebeca Niles és Michael Poston összegzi azokat a szerkesztési-kódolási tapasztalatokat, melyeket az illusztris Folger Shakespeare Library digitalizálási projektje, a Shakespeare összes művét kereshető, XML-fájlként is letölthető formában kínáló Folger Digital Texts készítése során szereztek. A digitalizálók feladata látszólag egyszerű volt, rendelkezésükre állt ugyanis a teljes korpuszt következetesen egységesítve közlő népszerű sorozat, a Folger Shakespeare Library szöveganyaga. Tanulmányuk érdekessége éppen abban rejlik, hogy egy viszonylag egyszerűnek tűnő példa, a szinte mindig szerkesztői beavatkozásra szoruló belső feliratok (beszélők jelzése, színpadi utasítások stb.) kódolása során felmerülő problémák sorával bizonyítják, hogy a kódolás során nagyon sokszor kell olyan döntéseket meghozni, melyek messze túlmutatnak az egyszerű kódoláson. A nagyon is gyakorlatias vállalkozás szövetében szervesen jelen vannak, és utalás szintjén a bevezető részekben is felidéződnek az utóbbi évtizedek fontos textológiai megfontolásai, McGann, McKenzie, McCleod és mások elméleti észrevételei, valamint a „szerkesztetlen” Shakespeare-t sürgető iskola (Leah Marcus) bizonyos megfontolásai is. Az elemzés során az adatmodell lesz az egyik kulcsfogalom, és különösen tanulságos, ahogy az ehhez hasonló, a számítástechnika világából származó fogalmakat visszavezetik a nyomtatott kiadásokra, melyekben, mint bemutatják, szintén mindig működik egyfajta implikált adatmodell – ez a fajta összekapcsolás sokat segít a digitális és a nemdigitális világ között esetlegesen érzékelt

konceptuális távolság csökkentésében. A tanulmány kellően nagy, de még kezelhető méretű, ugyanakkor kifejezetten népszerű szövegcsoporthoz, az elmélet és a gyakorlat ütköztetésével mutatja be, hogy a digitalizáció miatt megváltozott olvasói szokásokat és elvárásokat (például a távoli olvasás igényét) megfelelően kiszolgáló kiadások kódolói lényegében a szerkesztők szerepét töltik be, mindeközben a digitalizáció általánosabb kérdéseire is gondolatébresztő módon reflektál.

A kódolási vonalon haladunk tovább a következő tanulmánnyal, amely akár már a kötet korábbi pontján is előkerülhetett volna. Martin Mueller, Philip R. Burns és Craig A. Berry ugyanis az Early English Books Online – Text Creation Partnership (EEBO-TCP) nevű digitalizálási-átírási-kódolási projekt kódolási háttéréről számol be, ám eddigre az esetlegesen a lineáris olvasás kényszerét követő olvasó már több olyan tanulmányon túl lehet, amelyek eredményei pontosan erre a korpuszra építenek. Az EEBO-TCP keretében már most is elérhető 25000 szöveg, ám 2020-tól a teljes 66000, 1473 és 1700 között kiadott könyvet magába foglaló adatbázis (a szerzők számítása szerint ez az összes megjelent könyv legalább 92%-át jelenti) szabadon hozzáférhető lesz. Várakozásaik szerint ez ma még el sem képzelhető mértékben fogja megváltoztatni a korszak tanulmányozását, ezért is nagyon fontos, hogy megértsük, milyen kódolási-gondozási folyamatok eredményeképpen is jutnak el hozzánk az érintett szövegek. A végső soron kíváncsi modell ismertetése során a szerkezeti-nyelvészeti annotációt a könyvtári katalógusokhoz hasonlítják, melynek eredménye egy háromszintű ideális szerkezet: ebben a metaadatok képezik a legfelső szintet, középen a nyelvi tagolás szintje, míg legalsó szinten az egyes szóelőfordulások találhatók. Nem csupán arról beszélünk tehát, hogy hogyan lehetne 66000 szöveget megfelelő metaadatokkal ellátni, hanem ideális esetben a teljes korpusz mintegy kétmilliárd lexikális elemének a „mély szöveggondozását” (*deep curation*) tűzzük ki célul, melyek elképesztően összetett, a mai keresőmotorokkal még le sem kérdezhető lekérdezéseket tesznek lehetővé. A kora újkori szövegek esetében az egyes szavak írásmódjában mutatkozó nagyfokú szórás nagy kihívást jelent, és a tanulmány a nyelvészeti annotáció technikai háttérét járja körül, bemutatva, hogyan segítik elő különböző eszközök (*CLAWS*, *MorphAdorner*) a szótövesítést, a szófajok szerinti címkézést és a kollaboratív szöveggondozást és -javítást. Ennek egyik legfontosabb része a lexikális elem (*token*) kezdeti azonosítása és a későbbi gondozási folyamatok különválasztása. A meglehetősen mély technikai ismertető után a gondozás társadalmi vonzatáról olvashatunk, arról, hogy a ma még sok hiányt és hibát tartalmazó korpuszt hogyan lehetne a felhasználók bevonásával feljavítani, és ennek a szoftveres háttérét, a problémás lexikális annotációk javítását és a javítások nyomon követését segítő *AnnoLex* programot is ismerteti. Itt is felmerül persze egy tudományfinanszírozási/-politikai probléma: ha hibát észlel, a program megjeleníti az eredeti oldalt az EEBO adatbázisából, az viszont természetesen csak előfizetés ellenében használható – a szövegek ugyan 2020-ban nyilvánosak lesznek, de a szkennelt képek továbbra is a ProQuest tulajdonában maradnak. A dolgozat utolsó része további javaslatokat fogalmaz meg, miközben analógiák sorával mutatja be, hogy a gigászinak tűnő vállalkozás, a gyakorlatilag szavanként végrehajtandó annotálás valójában nem olyan ijesztő dolog, mint amilyennek elsőre hangzik.

A következő dolgozatban Anupam Basu a nyelvészet irányába fordul, és azt próbálja meg felülvizsgálni a digitális módszerek segítségével, hogy mennyire helytállóak az

angol nyelv helyesírásának sztenderdizációjával kapcsolatban általánosan elfogadottnak tekinthető szakirodalmi alaptételek. Ezek a 16–17. század fordulóján megjelenő elméleti szövegek fontos, de nehezen kimutatható szerepét emelik ki, míg a folyamatot úgy írják le, mint amely hosszan tartó, lassú fejlődést követően a 17. század közepére, a polgárháborúval párhuzamosan ért el az áttörés fordulópontjára. A rendkívül izgalmas módszertant alkalmazó dolgozat a korábbiakhoz hasonlóan az EEBO-TCP korpuszára épít, és abban figyeli meg néhány példán keresztül, hogy az egyes szavak alternatív írásmódjainak az időbeli gyakoriságát ábrázoló függvények hajlamosak hasonló mintázatot követni. Ezután a szerző alapos matematikai dokumentációval együtt mutatja be, hogy szavak helyett ilyen mintázatokat keresve a korpuszban hogyan tudta újragondolni a sztenderdizáció kérdését. A keresés egyértelműen feltárta, hogy az nem egy hosszan tartó, lineáris folyamatként, sokkal inkább különböző variánsok egymással vívott, az egyes szavak vagy pusztán grafémák szintjén újra meg újra lejátszódó, hullámszerűen terjedő evolúciós csatájaként kell elképzelni. Ily módon, bár a szakirodalom korábbi megállapításai az időt illetően helytállóak, a háttérben egy, az eddigtől homlokegyenest eltérő narratíva figyelhető meg, ami ismét csak a nagy, forradalmi változásokról szóló narratívákat látszik megkérdőjelezni.

Timothy W. Cole, Myung-Ja K. Han, Mara R. Wade dolgozatának középpontjában konkrétan az Emblematica Online emblémáskönyv-adatbázis, általánosabban viszont az összekapcsolt nyílt adatok (*linked open data, LOD*) és a szemantikus webes technológiák gyakorlati alkalmazhatósága áll. Voltaképpen nem is annyira dolgozattal állunk szemben, inkább egyfajta igen részletes műszaki dokumentációval, így némileg meglepő módon szinte csak érintőlegesen, felvillantott példákon keresztül olvashatunk arról, hogy a digitalizáció hogyan változtatta meg a kora újkori emblémakönyvek hagyományával kapcsolatos tudásunkat. Sokkal hangsúlyosabb a dolgozatban magának az adatbázisnak a kialakításában és fejlesztésében követett módszer, melynek az volt a lényege, hogy az adatbázist már a kezdetektől nyitott, fejleszthető, más adatbázisokkal együttműködni tudó (interoperabilitásra kész) szolgáltatásként tervezték. Az emblémák esetében az átjárást nagyban segítette az LOD-technológián keresztül az adatbázisban is alkalmazott *Iconclass* szolgáltatás, amely egyrészt az emblémák egységes leírásában, másrészt a későbbi, akár emblémakönyveken kívüli forrásokkal történő összekapcsolásban segít. Így végső soron arra a kérdésre is választ kapunk, hogy maga az emblémakutatás hogyan változhat mindettől: egyes, földrajzilag szétszórt emblémakönyvek helyett a belefoglaló kiadványtól függetlenített, tematikusan kapcsolódó emblémák hálózatszerű kutatása válik lehetővé, és a más szolgáltatásokhoz való LOD-kapcsolódás a páneurópai kontextus bevonásával szintén ezt a hálózatoságot hivatott gazdagítani.

A következő dolgozat mintha egy külön ki nem emelt utolsó szakaszt vezetne be a kötetben belül, a hátralévő tanulmányok nagy része ugyanis valamilyen módon kapcsolódik a tér fogalmához, a bölcsészettudományban is lezajlott térfordulathoz. Ez Janelle Jenstad és Diane K. Jakacki tanulmányára is igaz, ugyanakkor a helynevek kora újkori drámákban történő feltérképezéséről szóló dolgozat, hasonlóan az előző tanulmányhoz, az egymástól függetlenül gondozott adatbázisok összekapcsolásának a lehetőségeire is nagy hangsúlyt fektet. Bár középpontjában a Maps of Early Modern London projekt áll, azt az Internet Shakespeare Editions kiadási vállalkozással, azon

belül is a *Henry VIII.* szövegével összefüggésben ismerteti. A térképészeti vállalkozás érdekessége, hogy Agas 1560-as, nem referenciális London-térképét szövegkiadási és -címkézési eljárásokkal teszi gazdagabbá. Ehhez különböző, Londonhoz kapcsolódó korabeli szövegek mellett drámákat is felhasznál, és a dolgozat azokat a kihívásokat dokumentálja, amelyekkel a készítő az említett dráma kiadásának a térképészeti program címkéivel kombinált előállítása során szembesültek. Az implikált vagy éppen az időben változó/kicserélődő és más okból problémás helynevek esetében a kihívások nagy része abból fakadt, hogy a kétféle projekt más-más céllal közelítette meg ugyanazt a szöveget. A számtalan így születő újszerű kutatási kérdés mellett a legfontosabb tanulság az, hogy az interoperabilitás érdekében nem lehet előre rögzített elméleti keretekhez ragaszkodni, a kereteket sokkal inkább a gyakorlat során kell kialakítani. Ebben pedig kulcsfontosságú szerepe van a programozóknak, akik ebben a dolgozatban kiemelt figyelmet kapnak.

A térképezetek a közösségi médiával kombinálva jelennek meg Fabrizio Nevola tanulmányában. A szerző az urbanizáció elméleti és történeti vonatkozásait felvilágosítva arra keresi a választ, hogy hogyan segíthetnek a modern helymeghatározó (GPS) és az azokra épülő információs (GIS) technológiák olyan végtermékek, jelen esetben okostelefonon futtatható alkalmazások előállításában, amelyek a szélesebb tömegek érdeklődésére is számot tarthatnak. Az elméleti háttér és korábbi geobölcsészeti projektek ismertetését követően a *Hidden Florence* nevű alkalmazás fejlesztéséről és működéséről olvashatunk esettanulmányt. Az alkalmazás korabeli térképre vetítve, a felhasználó mozgását követve előre rögzített félig fiktív hanganyaggal mutatja be Firenze epizodikus (mikro)történetét. A vállalkozás talán legizgalmasabb eredménye, hogy miközben narratív-interaktív jellegével bevonja a nem feltétlen tudományos érdeklődésű felhasználót, történetmesélésében a mikrotörténetírás Levi- és Ginzburg-féle iskolájának a szempontjait is sikerrel érvényesíti, és így már-már a gamifikáció irányába mutatóan csempész rá tudományos eredményeket a tömegek okoskészülékeire.

A következő projekt már nem egyszerűen a térrel, hanem a virtuális térrel foglalkozik, sőt, John N. Wall dolgozata még ennél is tovább megy, és a kora újkori hangzás digitális rekonstrukciójának a lehetőségeit méri fel. A dolgozat középpontjában a londoni Szent Pál-székesegyház 1666 előtti változatának a digitális modellje áll, ám ezúttal nem csupán a londoni tűzvészben elpusztult épület háromdimenziós vizuális modelljéről van szó, hanem a prédikációkutatás kérdéseitől hajtva, azokat továbbgondolva a *Virtual Paul's Cross* projekt keretében az épület hangmodellje is elkészült. Talán az egész kötetben ez a legfuturisztikusabb vállalkozás, amelynek hátterét az élvonalbeli műszaki technológiák és a kora újkorral kapcsolatos kutatási eredmények hatékony elegye képezi. Mindaz alapján, amit történeti forrásainkból a székesegyházról, az annak udvarán tartott, nagy tömegeket megmozgató prédikációkról, azon belül esettanulmányként John Donne egyik 1622-es prédikációjáról tudunk, professzionális modellezők és hangmérnökök segítségével sikerült reprodukálni azt, hogyan hangozhatott az adott prédikáció eredeti közegében. Az új megközelítés persze nehéz kérdéseket is felvet, hiszen az így kapott virtuális prédikáció nyilvánvalóan soha nem hangzott el. Ugyanakkor a modellezésből egyértelműen és vélhetőleg megbízhatóan derül ki, hogy az itt elhangzott prédikációkat az óránkénti harangszó miatt meg kellett

szakítani, ami alighanem a megkomponálásukra is eleve hatott – vagyis egy lépéssel közelebb kerülhetünk így a prédikáció valódi közegének a megismeréséhez. A dolgozat arra hívja fel a figyelmet, hogy eljött az idő, amikor a természettudományokhoz hasonlóan a bölcsészettudományban is lehet már bizonyos jelenségeket modellezni – ám az így kapott eredmények megítélése jelenleg még igencsak bizonytalan.

A következő dolgozat, Jennife Robert-Smith, Shawn DeSouza-Coelho és Paul J. Stoesser tanulmánya hasonló irányban vizsgálódik, de a *Simulated Environment for Theatre* vállalkozás örvén a színháztörténet területére kalauzolja az olvasót. Ahogy a kötet előrehalad, egyre naprakészebb technológiák kerülnek elő, így már szinte meg sem lepődünk, hogy az önmagát drámák olvasására, felfedezésére és rendezésére szánt 3D-s vizualizációs környezetként definiáló SET a *Unity* játékmotorra épül, és a szerzők egyike a videojáték-kutatás területéről érkezett. A dolgozatban azt igyekeznek kideríteni, hogyan segítheti a rendszer a forrásalapú színházkutatást, és az esettanulmány tárgyaként az 1540 és 1640 között szezonálisan működő cambridge-i Queen's College színház szolgál, ám leszögezik, hogy a fő cél ennél általánosabb: kideríteni, hogyan alkalmazhatók a digitális eszközök, különösen a szimulációk a színháztörténet területén. A tanulmány egyedi hangnemet üt meg, ugyanis meglepő módon nem egyértelműen pozitív a végkövetkeztetés. Sőt, mintha az bontakozna ki, hogy a színház szerintük legfontosabb jellemzője, annak fenomenológiája egyáltalán nem reprodukálható a szimuláció segítségével. Végző soron azonban arra jutnak a szerzők, hogy ez nem feltétlen probléma – valahogy úgy van ez, mint ahogy a történeti irányultságú kutatásokkal általában: nem elegendő mennyiségű adatból voltaképpen a történész is szimulál egy történetet, ami nyilvánvalóan nem azonos a valós eseménnyel, ám a szimuláció önmagában is hasznos, abból önmagából is tanulunk valamit (mintha Huizinga játék-tudomány párhuzama sejlene fel a háttérben). A dolgozat arra a következtetésre jut, hogy a felmerült problémák végző soron a teatralitással, sőt magának a színháztörténetnek mint tudománynak a teatralitásával kapcsolatban is sok új megfontolást szültek, így hasznuk megkérdőjelezhetetlen.

És ha a videojáték-kutatás után valaki netán még hiányolta volna a közösségi hálózatokat a kötet egyébként rendkívül széles merítéséből, annak Geoffrey Way zárótanulmánya hoz megnyugvást. A tanulmány talán azért is kerülhetett a kötet végére, mert bizonyos tekintetben már kifelé mutat. Egyáltalán nincs ugyanis itt szó kora újkori drámatörténetről, jószerivel a kora újkorról is csak érintőlegesen. A szerző elsősorban abban érdekelt, hogy különböző, Shakespeare-hez kötődő mai színházak hogyan próbálnak meg túlélni az egyre szűkülő gazdasági lehetőségek közepette a digitális marketing, különösen a közösségi oldalak segítségével. Bemutatja, hogyan igyekeznek, általában a nosztalgiára alapozva interaktív kapcsolatot létesíteni a közönségükkel, ami több szempontból is előnyös. A tényleges látogatók újabb látogatásra serkenthetők az elköteleződés növelése révén, ám a csak online jelen lévő közönség is üdvös, amennyiben erősíti az adott intézmény kulturális jelentőségének az érzékelhetőségét. Mindazonáltal a kissé hideg marketinges megfontolásokat tartalmazó, a kötet tárgyköréből némileg kilógó tanulmány meglátása szerint, bár a vizsgált kezdeményezések jó irányba mutatnak, az intézmények előtt még akad tennivaló, ha igazán versenyképesek akarnak lenni a „figyelemgazdaság” területén.

A fentiekből már alighanem kiderült, hogy a kötet rendkívül, már-már túlradóan tartalmas és sokszínű. Annyiféle témát érint és azokat annyiféle szempontból dolgozza fel, hogy az szinte túlfeszíti egy kötet kereteit. De talán éppen ez a bőség és sokféleség kell ahhoz, hogy beláthassuk: a digitális fordulat már valóban bekövetkezett, és a kora újkor kutatásának a legkülönbözőbb területeit forgatta fel. Érdekes észrevenni, hogy a számadás mennyire megerősíti Leah Marcus két évtizeddel ezelőtti jóslatait. A neves kutató 1995-ös tanulmányában („Cyberspace renaissance”) ugyanis azt jósolta, hogy az éles vonalak, a konfliktusok, forradalmi fordulatok narratívája az, ami a digitális fordulatot követően el fog tűnni. A kötetben ez az érzés újra meg újra felerősödik, a nagy korpuszon végzett, automatizált vizsgálatok egyre természetesebbnek láttatják az irodalomtörténeti, nyelvészeti stb. változások folyamatjellegét, vagyis a digitális módszerek valamelyest mintha neutralizálni tudnák egyrészt a nagy narratívák hatását, másrészt a kánon szelekciós természetéből fakadó problémákat.

Mindeme változások horderejét Marcus annak idején a kéziratosságról a nyomtatásra való átállás fontosságához hasonlította. A vizsgált kötet szintén megerősíti a digitális fordulat óriási horderejéről vallott nézeteket, ugyanakkor a kora újkori váltáshoz hasonlóan itt is megfigyelhető az újféle médium kezdeti bizonytalansága. A kötetet forgatva ugyanis az olvasó (vagy inkább felhasználó?) újra meg újra azt érzi, hogy ez az esszégyűjtemény hatékonyabban működhetne digitális formában. Ennek eklatáns példái a kötetben nagy számban található, sok esetben kulcsfontosságú illusztrációk – azok ugyanis a nyomtatott változatban alig bogarászhatók ki, esetenként teljességgel használhatatlanok. Ezért nem csupán kényelmes, de egyenesen nélkülözhetetlen a könyvhöz készített, szabadon elérhető segédhonlap,¹ melynek hiánya jelentős mértékben megnehezíti a befogadást. Ezt a jelen ismertető írója akkor érezhette át igazán, amikor ismeretlen okból egyik napról a másikra már nem működött a honlap, és emiatt életében először hibajelentést kellett küldenie nem egy szoftver, hanem egy nyomtatott könyv kiadójának. Példás gyorsasággal reagáltak, a honlap a megírás pillanatában már újra működik. Ugyanakkor a korai nyomtatott könyvekre jellemző gyermekbetegségek még így is előfordulnak: az egyik tanulmány a honlapon bizonyító érvényű videós modellt ígér, ám az nem szerepel ott, amittől közvetlenül a dolgozat argumentációja akad meg, máskor meg egyszerűen csak azon gondolkodik el az ember, hogy miért kell ezen az anyagon két médiumnak osztoznia, vagy hogy mi keresni valója van a nyomtatott kiadványban a képernyőn alkalmazható kiemelések hiányában nehezen áttekinthető XML-részletnek, hogy nem idejétmúlt-e egy, a hipertextről szóló szövegnek a végén elhelyezni egy szószeretet – és ez a lista korántsem teljes.

Mégis, e nyomtatott-digitális hibridről összességében kijelenthető, hogy kitartó olvasója nem pusztán néha már-már kíméletlen alaposságú, szerteágazó, nagyon sokféle technológiát felvonultató tömbösített kurzust teljesít digitális bölcsészethől, hanem a folyamatos önreflexiónak, a szinte mindenhol megfigyelhető, a tágabb kutatási területet érintő érzékeny problémafelvetéseknek köszönhetően a kora újkor kutatását érintő legalapvetőbb kérdések újragondolására is késztetést kap. A kötet folyamatosan érdekfeszítő jellege mellett sem könnyű olvasmány, kivételesen sok utánanézést

¹ „Early Modern Studies After the Digital Turn,” hozzáférés: 2018.06.08, <https://ems.itercommunity.org/>.

igényel, ám az olvasó e fáradságért cserébe rengeteg továbbgondolandó kérdéssel, kutatási ötlettel, és ami a legfontosabb: felszabadító inspirációval gazdagodik, legyen bár az az olvasó az informatika iránt érdeklődő bölcész, vagy a bölcsészet iránt érdeklődő informatikus. És a kötet azt is sugallja, hogy e két alak a jövőben várhatóan egyre kevésbé lesz szétválasztható.

Zámbóné Kocic Larisa

Szegedi Tudományegyetem, Angol Tanszék

larisa@ieas-szeged.hu

Matthew K. Gold, ed. *Debates in the Digital Humanities*. Minneapolis–London: University of Minnesota Press, 2012. ISBN 9780816677948, 9780816677955. xvi, 516 oldal. <https://doi.org/10.5749/minnesota/9780816677948.001.0001>

Matthew K. Gold and Lauren F. Klein, eds. *Debates in the Digital Humanities 2016*. Minneapolis–London: University of Minnesota Press, 2016. ISBN 9780816699537, 9780816699544. xv, 579 oldal. Hálózati verzió: <http://dhdebates.gc.cuny.edu/debates/2>

Az alábbiakban a University of Minnesota Press gondozásában 2012-ben és 2016-ban megjelent *Debates in the Digital Humanities* [Viták a digitális bölcsészetben] két kötetéről igyekszem röviden összehasonlító recenziót írni. A Matthew K. Gold szerkesztette első kötet 2012-ben azzal a céllal jött létre, hogy egy kötetbe gyűjtse az amerikai felsőoktatásban a „következő nagy dobás”-ként („next big thing”)¹ emlegetett, de igazán pontos meghatározással még nem bíró, digitális bölcsészetről szóló vitákat. Az első kötet igazi értéke véleményem szerint abban rejlik, hogy a meghatározási vitákon túl – amelyek a kötet első részét alkotják² – a digitális bölcsészettel foglalkozó korábbi vagy épp kortárs könyvekből³ hiányzó kritikai hangoknak is helyet ad. Különösen két téren fontos ez a hiánypótlás: a digitális bölcsészet tanításának, illetve a kritikai kultúrakutatásnak kérdéskörében. Ez utóbbi kapcsán írja Alan Liu, hogy noha

[...] a digitális bölcsészek kritikusan fejlesztik az eszközöket, az adatokat és a metaadatokat [...] a kritikájukat ritkán terjesztik ki a társadalom, a gazdaság, a politika vagy a kultúra teljes regiszterére. Ritkán hallani az általam ismert digitális bölcsészeti társaságokban, konferenciákon, folyóiratokban vagy projektekben olyan kérdéseket, amelyek azt vizsgálják, hogy a digitális bölcsészet

¹ William Pannapacker, „Pannapacker at MLA: Digital Humanities Triumphant?” *The Chronicle of Higher Education*, 2011. jan. 8., <https://www.chronicle.com/blogs/brainstorm/pannapacker-at-mla-digital-humanities-triumphant/30915>. E blogbejegyzés később szerves része lett a Gold szerkesztette *Debates in the Digital Humanities* (2012) gyűjteménynek.

² Többek között Matthew Kirschenbaum, Kathleen Fitzpatrick és Lisa Spiro írásaiból.

³ Susan Schreibman, Ray Siemens and John Unsworth, eds., *A Companion to Digital Humanities* (Malden, MA: Blackwell, 2004); David M. Berry, ed., *Understanding Digital Humanities* (Basingstoke: Palgrave Macmillan, 2012); Melissa Terras, Julianne Nyhan and Edward Vanhoutte, eds., *Defining Digital Humanities: A Reader* (Farnham and Burlington: Ashgate, 2013).

hogyan szorgalmazza, képviseli napjaink hatalmas posztindusztriális, neoliberais, nagyvállalati és globális (információ)tőkéjét,⁴ vagy áll éppen ellen neki. E kérdések klisészerű megnyilvánulásai – mint például a „digitális szakadék”, „megfigyelés”, „magánjog”, „szerzői jog” stb. – sem rúgnak nagyon a labdába (491).⁵

Persze az olyan kérdések felvetése, hogy miért „fehér” a digitális bölcsészet (139–160), illetve, hogy van-e mindebben szerepe a nőknek (235–240), visszacsatol a digitális bölcsészet legmarkánsabb identitás vitájába. Vajon csak az a digitális bölcsész, aki *homo faber*, aki digitális adatbázisokat és azokat elemző algoritmusokat hoz létre, vagy ide sorolható a *homo logicus* is, aki a bölcsészet mai eszköztárával igyekszik értelmezni digitális korunk megnyilvánulásait (többek között a fent említett kérdésekre is választ keresve)?⁶ Erre a kérdésköre még visszatérek a recenzióm vége előtt. Annyit azonban már itt elmondhatok, hogy e válaszra aligha kapunk még egy ideig kielégítő választ, annál is inkább mert – ahogy David M. Berry és Anders Fagerjord tavaly megjelent könyvében olvassuk – immáron átléptük a posztdigitális kor küszöbét, amelyben már nem a digitális és az analóg kultúra egymásnak feszülése a kérdés, hanem a digitális megnyilatkozások átfogó volta, és az ez okozta változások, amelyek fényében szükségszerű lassítani a folyamatokon és megkérdőjelezni azokat a további digitális építkezés mellett is.⁷

A *Debates in the Digital Humanities 2016* kötet lényegében a fenti hagyományt követi, egy kötetbe összesítve a digitális bölcsészetről azóta megjelent jelentősebb, vitaindító tanulmányokat. A szerkesztők, Matthew K. Gold és Lauren F. Klein, öt részbe sorolták e tanulmányokat, attól függően, hogy a DH múltjaira és jövőire, gyakorlataira, diszciplináris megnyilvánulásaira, bírálatára, vagy a digitális szövegelemzés lehetőségeire összpontosítanak. Az előző kötethez hasonlóan a tanulmányok mellett itt is minden könyvrészt kettő vagy három rövid esszéírás és/vagy blogbejegyzés zár. Gold és Klein szerint ugyanis a blogoszférában és a Twitterverzumban megjelenő „szürke irodalom” központi szerepet játszik a digitális bölcsészet artikulációjában. Mi több, a digitális bölcsészetről szóló viták szerteágazó platformokon való megnyilvánulása teszi szükségszerűvé a vitagyűjtemény (tervek szerinti) újabb kötetének évről évre való megjelenését (569).

A korábban hiánypótló témaként külön részt kapó tanítási gyakorlat most szervesen épül a fent említett részekbe, mindazonáltal külön figyelmet érdemel talán Ryan Cordell tanulmánya arról, hogy hogyan ne tanítsuk a digitális bölcsészetet. Cordell konkrétan azt javasolja, hogy a szakirodalom jelentős részét képező meghatározási vitákat úgy ahogy vannak hanyagoljuk, mivel „túl gyakran elmulasztjuk

⁴ Az angol eredetiben szereplő *information-cum-capital* kifejezés arra utal, hogy a szóban forgó nagyvállalatok tőkeként kezelik az információt, de egyben valós tőkét is kovácsolnak belőle. Lásd a napjainkban zajló Facebook–Cambridge Analytica botrányt.

⁵ A recenzióban az angol eredetiből magyarul idézett szövegrészek saját fordítások.

⁶ Lásd erről bővebben Kathleen Fitzpatrick, „The Humanities, Done Digitally,” *The Chronicle of Higher Education*, 2011. máj. 8., <https://www.chronicle.com/article/The-Humanities-Done-Digitally/127382>. Ez a rovatcikk is része lett a DH-viták első kötetének.

⁷ David M. Berry and Anders Fagerjord, *Digital Humanities: Knowledge and Critique in a Digital Age* (Cambridge, England–Malden, MA: Polity Press, 2017), 2.

bizonyítékokkal alátámasztani a változásokat hirdető állításainkat” (462). Ez utóbbira jó példa a módszertani részben olvasható Alex Gil-szöveg („Interview with Ernesto Oroza”), amely többek között egy ígéretes felvetéssel kecsegtet a magyar bölcsész olvasó számára: „Hogyan dolgozhat valaki számítógépekkel a bölcsészetben, ha azt egy olyan közegben kell gyakorolnia, ahol a géppark több mint tíz éves, az internet elérhetőség pedig bizonytalan vagy nem létező?” (184). Sajnos, bármennyire is szeretnénk, a szövegből erre nem kapunk választ, mert a szerző interjúja Ernesto Oroza kubai építésszel kimerül abban, hogy inspirációt lát az Oroza által is képviselt kubai *csináld-magad* (DIY) kultúrában, amelynek szemléletmódját Gil szerint alkalmazni kellene/lehetne a digitális infrastruktúra szempontjából kihívások elé állított helyeken, azonban nem tér ki rá, hogy lehetne ezt konkrétan megvalósítani, és így megreked a lehetőségek, a potenciák üres hangoztatásában. Pozitívabb példaként itt érdemes talán megemlíteni Miriam Posner tanulmányát (32–41) – a digitális bölcsészet múltjait és jövőit latolgató részből – amely címében ugyan maga sem kerüli el a lehetőségek túlexponálását: „What’s Next: The Radical, Unrealized Potential of Digital Humanities [Mi jön ezután: A digitális bölcsészet radikális, (eddig) meg nem valósított potenciája].” Ennek ellenére Posner konkrétan rávilágít az üzleti alkalmazásokból kölcsönvett adatfeldolgozási (vizualizációs) modellek egysíkúságára, amelyek remekül alkalmazhatók konkrét adatok megjelenítésére, ám kevésbé jók, ha összetettebb vagy bizonytalan adatokat kell szemléltetni (például egy olyan idővonal esetében, ahol nem áll módunkban percre, másodpercre pontos adatokat szolgáltatni, mert mindössze annyi információ áll a rendelkezésünkre, hogy „kb. az 1500-as években,” 33). Ezek a kölcsön- és/vagy átvett modellek olyan hatalmi struktúrákkal (pl. faj, társadalmi nem) dolgoznak, amelyek leegyszerűsített sémáik miatt nem mennének át a bölcsészettudomány rigorózus szűrőin, ezért Posner szerint el kell gondolkodnunk azon, hogy hogyan építhetnék újra az ilyen reduktív logikával dolgozó, archívumokat és adatbázisokat fenntartó és működtető programokat, algoritmusokat. Cordellhez hasonlóan Posner is kitér a DH, szerinte céltévesztett oktatására, amikor azt állítja, hogy „a digitális bölcsészet egyetemi oktatásának értéke nem abban rejlik, hogy új technológiák alkalmazására tanítjuk a hallgatókat, hanem hogy megmutathatjuk nekik, mennyire ideiglenes, relatív és ideológiailag átítatott az adatok által felépített és bennünket körülvevő világ egésze” (35). Szerencsére a kritika mellett Posner pozitív példákat is említ, amelyek összetettebb adatfeldolgozással és szemléltetéssel (infrastruktúrával) kísérleteznek (36–38).

Három tanulmány különösen felkeltette irodalmár érdeklődésem: Jeffrey M. Binder cikke („Alien Reading: Text Mining, Language Standardization, and the Humanities,” 201–219), amely rámutat a témamodellezés (*topic modelling*) sztenderd nyelvhasználat iránti elfogultságára, és ezáltal az irodalmi szövegekre való alkalmazhatóságának korlátjaira; valamint Michael Hancher és David L. Hoover szövegei, amelyekben a szerzők egymástól függetlenül is a digitális bölcsészet által felturbózott szövegközeli avagy szoros olvasás (*close reading*) fontosságát hangsúlyozzák, más-más megközelítésből. Hancher („Re: search and Close Reading,” 118–138) abból indul ki, hogy ideje a digitális bölcsészeten belül mostanra közhelynek számító makro szint (adatok kinyerése és leképezése) jelentősége mellett a mikro szinten (szinguláris adat kinyerés) történő változásokat és lehetőségeket, valamint a Franco Moretti által népszerűsített távoli

olvasás (*distant reading*)⁸ mellett a szoros olvasást is kiemelni a digitális bölcsészet vívmányai között. Hancher szerint ugyanis

[...] a távoli olvasás nem a szoros olvasás ellentéte. A szoros olvasás szövegközelisége régóta nem a térbeli közelségről szól [ahogy a távoli sem a térbeli távolságról], hanem a sűrűségről és a koncentrációról: az olvasó koncentrációjáról mindenképpen, ám gyakran az olvasott szöveg koncentrált voltáról, sűrűségéről is. Ezt jelentette az új kritikusok és a régiek számára egyaránt. A szoros olvasás ellentéte nem a távoli olvasás, hanem a laza, az esetleges és a figyelmetlen olvasás (125).

David L. Hoover az „Argument, Evidence, and the Limits of Digital Literary Studies [Érvelés, bizonyíték és a digitális irodalomtudomány korlátjai]” (230–250) című tanulmányában Hancherhez hasonlóan a szövegközelis olvasásban látja a digitális bölcsészet elmúlt huszonöt év felhozatalának legfőbb értékét, és az ezt alátámasztó két tanulmány állításait vizsgálja újra digitális módszerek segítségével, hogy akkurátusabb, a szöveg belső logikájával összeegyeztethető konklúziókra jusson. Meglátása szerint ugyanis a digitális bölcsészet legfőbb értéke abban rejlik, hogy feltárhatja az irodalomtudományban tévesen használt, vagy legalábbis félreértelmezhető érvrendszereket és bizonyíték alkalmazásokat, amelyek lehetővé tették az 1990-es évek végén elhíresült Sokal-botrányt.⁹

Érdekes megfigyelni, hogy bár a kötet külön részt szentel a DH-kritikáknak, a többi részben összegyűjtött szövegek éppúgy hangot adnak kritikai hangoknak, és egy-két túlkapástól eltekintve egyikre sem jellemző a korai kiadványok tartózkodás nélküli DH-kampánya. Csakugyan beigazolódni látszik Berry és Fagerjord korábban már említett megállapítása, miszerint a DH immáron egy olyan korba lépet, amelyben korábbi pörgő lendülete elkerülhetetlenül alábbhagy és már csak óvatosabb léptekkel építkezik tovább.

De ha ennyire szervesen jelen van a kritika és az önreflexió, akkor felmerülhet a kérdés, hogy minek egy külön részt szentelni a bíráló hangoknak? Úgy vélem, hogy a bírálatok a korábbiakhoz képest sokkal inkább politikai vonatkozásúak. Különösen igaz ez a „The Dark Side of Digital Humanities [A digitális bölcsészet sötét oldala]” című, társszerzős tanulmányra, amely a 2013-as bosztoni MLA konferencia azonos nevű kerekasztal beszélgetésének vitaindító előadásait hozza le, és amelyben Grusin

⁸ A Moretti-féle távoli olvasásról magyarul lásd Labádi Gergely, „Franco Moretti: *Distant Reading*,” *Irodalomtörténet* 95, 4. sz. (2014): 561–64; illetve annak magyar vonatkozású alkalmazásáról Péter Róbert, „Digitális és módszertani fordulat a sajtókutatásban: A 17–18. századi magyar vonatkozású angol újságcikkek »távolságtartó olvasása«,” *Aetas* 30, 1. sz. (2015): 5–30.

⁹ Alan Sokal fizikus 1996-ban leadott egy kamu tanulmányt a Duke egyetemi kiadó által gondozott *Social Text* kultúratudományi szakfolyóiratba, mintegy kísérletként, hogy egy szakmailag elismert folyóiratban megjelenhet-e egy olyan tanulmány, amely lényegében tartalmilag értelmetlen zagyvaság, de a) a szakzsargont alkalmazva jól hangzik és b) a kiadó ideológiai elvárásainak megfelelő álláspontot képvisel. A tanulmány megjelenését követően Sokal leleplező tanulmányt közölt a kísérletéről. A Sokal-ügyről részletesen Kampis György, az ELTE informatikai tudományokkal (tudomány- és bibliometria, adatbányászat, big data) is foglalkozó oktatójának honlapján lehet többet megtudni (lásd Kampis György, „A Sokal-ügy dokumentumai,” *Hps.elte.hu*, 2008, Hozzáférés: 2018.04.10, <http://hps.elte.hu/~gk/Sokal/Sokal.html>).

és társai a digitális bölcsészetben nem a hagyományos bölcsészet megmentőjét látják, hanem a társadalmi értékvesztésének egyik jelentős eszközét, amennyiben a digitális bölcsészet az állam által támogatott, profitorientált, a versenyszférában konkrétan hasznosítható ismeretek továbbadásának retorikáját erősíti.¹⁰ Sőt, az amerikai politikai szféra 2016 óta egyre inkább polarizálódó légköre a digitális bölcsészet vitáiba is betüremkedik azáltal, hogy a korábban már említett szemben álló felekként a *homo faber* kontra *homo logicus* helyett immáron a neoliberális kontra liberális kutató posztulálódik.¹¹ Ki tudja? Talán az épp ezen elmérgesedő vita miatt nem jelent meg újabb vitakötet 2017-ben.

Összességében véve a viták 2016-os kötete nagyon hasznos, mert valóban egy csokorba gyűjti a digitális bölcsészet legmarkánsabb vitáit és irányzatait, amelyből akár szemezgethetünk is a szakirodalom összegző szövegeinek magyarrá való fordításán gondolkodva.

¹⁰ Itt érdemes megjegyezni, hogy Pannapacker beszámolója a kerekasztal beszélgetésről a visszhangok tekintetében nagyon fontos, mert látványosan alátámasztja, hogy a kerekasztal résztvevői és az arról nyilatkozó kutatók merőben másként látják és értelmezik a digitális bölcsészetet. Ryan Cordell szerint a kerekasztal beszélgetés teljes félreértelmezése volt a digitális bölcsészetnek; Amanda French kijelentette, hogy egy digitális bölcsészt sem ismer, aki szeretné a MOOC-ot; Matthew Kirschenbaum pedig – nagyon helyesen – tiltakozott az elmélet és a gyakorlat szembeállításával szemben, mert, ahogy nyilatkozott: „csinálni [’fabrico’] annyi, mint gyakorlatba helyezni az elméletet: [a kettő szembeállítás] hamis dichotómia.” Részletesen lásd William Pannapacker, „On »The Dark Side of the Digital Humanities«,” *The Chronicle of Higher Education*, 2013. jan. 5., <https://www.chronicle.com/blogs/conversation/2013/01/05/on-the-dark-side-of-the-digital-humanities/>.

¹¹ E téren különösen nagy vitát generált az alábbi tanulmány: Daneil Allington, Sarah Brouillette and David Golumbia, „Neoliberal Tools (and Archives): A Political History of Digital Humanities,” *Los Angeles Review of Books*, 2016. máj. 1., <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.

Molnár Sándor Gyula

Szegedi Tudományegyetem, Klebelsberg Könyvtár

sandor.molnar@ek.szte.hu

George Bruseker, László Kovács and Franco Niccolucci, eds. „Digital Humanities.” Special issue, *ERCIM News* 111 (2017 október). 64 oldal.
<https://ercim-news.ercim.eu/en111>

Az *ERCIM News* 2017. októberi tematikus száma hasznos olvasmány lehet a digitális bölcsészeti kutatások aktuális trendjei iránt érdeklődő közösségeknek. Az ERCIM rövidítés az European Research Consortium for Informatics and Mathematics [Informatikai és Matematikai Európai Kutatási Konzorcium] nevet takarja,¹ amely egy európai gyökerű, nemzetközi közösséget jelöl. A szervezetet 1989-ben alapították, jelenleg tizenöt vezető kutatóintézetet foglal magába Európa tizennégy különböző országából. Hazai vonatkozásban talán az MTA SZTAKI áll jellegében a legközelebb az ERCIM-hez, és a két szervezet egymással kapcsolatban is áll. Központja Franciaországban található, egy Nizzától 25 km-re található technológiai parkban, a Sophia Antipolisban. A település voltaképpen egy egyetemi campushoz, illetve egy kutatóintézeti parkhoz hasonlatos. Itt található többek között a W3C európai központja, és számos egyéb informatikával, elektronikával, biotechnológiával és gyógyszeriparral kapcsolatos kutatóintézet is. Magyarországon hasonló – talán kisebb és kevésbé összetett – központ a debreceni Science Park, illetve Szegeden, a fejlesztés alatt álló ELI Science Park. Az ERCIM célja előmozdítani az európai kutatóközösség együttműködését és kapcsolódásukat az európai ipari vállalatokhoz. Tevékenységéhez tartozik tudományos publikációk és hírlevelek kiadása, workshopok, szemináriumok és posztdoktori képzések szervezése, valamint olyan nagy volumenű európai pályázatok, mint például a HORIZON 2020 menedzselése.² Maga a folyóirat, az *ERCIM News* negyedévente jelenik meg, hozzávetőlegesen 6000 nyomtatott és roppant elegáns példányban, de online is elérhetőek a lapszámok. A lap archívuma visszamenőlegesen tartalmazza a korábban megjelent összes számot. A folyóíratra jellemző, hogy az egyes számok egy-egy adott témával kapcsolatos írásokat tartalmaznak. Ezek a tárgykörök széles spektrumot fednek le, technológiai és elméleti kérdéseket egyaránt bemutathatnak, mint például az önvezérlő járművek, vagy a tudományos adatok megosztásának és újrafelhasználásának lehetőségei. A 2017 októberében megjelent tematikus szám különösen érdekes számunkra, hiszen digitális bölcsészet tárgykörébe tartozó írásokat gyűjtött össze. Rövid (jellemzően egy-két oldalas), gyakran több szerző által jegyzett cikkekből épül fel a szám. Ilyen módon az írások inkább a híradás igényével, figyelemfelkeltő jelleggel, mintsem részletes projektismertetésként íródtak.

¹ European Research Consortium for Informatics and Mathematics, hozzáférés: 2018.01.15, <https://www.ercim.eu/>.

² HORIZON 2020: The EU Framework Programme for Research and Innovation, hozzáférés: 2018.01.15, <https://ec.europa.eu/programmes/horizon2020/>.

A digitális bölcsészet témakört öt altémára bontották a szerkesztők, mi is e csoportosítást követve mutatjuk be a lapszámot. Az első alkategória a digitális források indexelésének és elemzésének kérdésköre. Ebben olvashatunk a Vatikáni Titkos Levéltár kéziratok gyűjteménye kapcsán a szövegfelismerési problémákról, illetve a KA3 nevet viselő projektről is, amelynek keretében szóban elhangzott szövegek automatikus feldolgozásának lehetőségeivel foglalkoznak. Tartalomelemzési szempontból izgalmas koncepció Ioana Manolescu cikke, amely a különböző sajtóorgánumban közölt állítások igazságtartalmának ellenőrzését, és ennek automatizálási lehetőségeit vizsgálja, amely különösen hasznos lehet az újságírók számára.

A második téma az információmenedzsment, illetve ennek stratégiai kérdései, melyben szövegekkel, illetve azok feldolgozásával kapcsolatos technikákat szemlélnek. A HumaReC-projekt technológiailag szövegátírási projekt, de egyfajta „folyamatos adatkiadási modellt” (*continuous data publishing*) vázol fel, mely a humán tudományok számára új típusú publikálási gyakorlattá válhat.

A harmadik tematikus egység a kutatásokat segítő infrastruktúra és eszközkészlet fejlesztéséről szól. Érdemes kiemelni a DIGILAB-ot mint az E-RIHS (European Research Infrastructure for Heritage Science)³ kezdeményezését, mely a kulturális örökség elektronikus feldolgozásával kapcsolatban hasznos eszköz lehet a jövőben. A DIGILAB célja, a feldolgozást segítő informatikai megoldásokat egy olyan platformba szervezni, mellyel a már meglévő adatbázisok tartalmát lehetne mélyebb szinten feldolgozni. Sajnos a cikk még nem mutat be új technikákat, a projekttel kapcsolatos konkrét szolgáltatások megjelenése majd csak az elkövetkezendő években várható. A negyedik témakör a háromdimenziós ábrázolásokkal, leképezésekkel foglalkozik. A bemutatott grafikai megoldások közül fontos említeni a középkori londoni Szent Pál-székesegyház digitális rekonstrukcióját, mely a hírhedt tűzvész előtti gótikus épületet mutatja be. A projekt célja az volt, hogy a napjainkra megsemmisült épületet minél részletesebben rekonstruálja digitálisan, a fennmaradt történeti források alapján.⁴ A cikk érdekes esettanulmány, de hozzá kell tenni, hogy technológiailag korszerűbb megoldásokat tartalmaznak a Pazirik Informatikai Kft. által készített termékek (rekonstrukciók és egyéb látványelemek) már hazai viszonylatban is. A Pazirik Kft. indulásakor történelmi 3D rekonstrukciókat készített, mára mindezt továbbgondolva mobil applikációkat, 3D nyomtatással készült „tapintható” múzeumpedagógiai eszközöket, valamint virtuális valóság alapú alkalmazásokat is fejlesztenek. Egy ezt követő rövid cikk („Culture 3D Cloud”) a háromdimenziós digitalizálás, és a modellalkotás folyamán fellépő megnövekedett hardverigény megoldására tesz javaslatot. Az ötödik tematikus egység az információvizualizáció és kommunikáció. Ebben a részben kerül bemutatásra a SZTAKI informatikai támogatásával működő Courage-projekt, amely a kelet-európai, volt szocialista országok területén lezajlott kulturális ellenállási törekvéseket gyűjti adatbázisba. Az események és a szereplők a *linked data* technológiát felhasználva az egymással kapcsolatban álló eseményeket szemantikus térbe helyezik.

³ European Research Infrastructure for Heritage Science, hozzáférés: 2018.01.15, <http://www.e-rihs.eu/>.

⁴ Virtual St Paul's Cathedral Project, North Carolina State University, hozzáférés: 2018.01.15, <https://vpccp.chass.ncsu.edu/>.

Az *ERCIM News* digitális bölcsészettel foglalkozó tematikus száma hűen ábrázolja e szakterület szerteágazó és sokrétű voltát. A lap és a cikkek tárgyalási módja alapvetően technológiai jellegű, de a bölcsészeti megközelítésű érdeklődő közönség számára is érdekes és főként informatív olvasmány lehet, hiszen számos olyan projekttel, módszerrel találkozhatunk benne, mely már akár a problémafelvetés szintjén is inspiráló és hasznos lehet.

<IN MEMORIAM>

Búcsú Labádi Gergelytől

Ravatal fölött állni és búcsúzni valakitől, akit szerettünk, de aki nincs többé, csak nyomok az üressé várt térben és az emlékezet, mindig nagyon nehéz. Labádi Gergelytől búcsúzni, leküzdeni a fájdalmat, és mondatokat formálni, mondatokba formálni a fájdalmat, nekem majdnem meghaladja az erőimet.

Mert fájdalmas, ha tanítvány búcsúztatja mesterét, de hatványozott a fájdalom és gyűlik a mélyről jövő, tehetetlen, fojtogató indulat, a lázadt ember vad keservei, ha a tanár búcsúztatja az egykori tanítványt.

Mert ez elfogadhatatlanul és elemien igazságtalan.

Másodéves kora óta ismertem őt. Máig emlékszem arra a régi szeptember eleji szemináriumra ott az Irodalmi Olvasóban, talán 1995 ősze lehetett, amikor először találkoztam velük. Ismerkedésképpen valami olyasféle feladatot adtam, hogy mindenki írjon pár sort a legutóbbi kortárs magyar irodalmi élményéről, arról, ami utoljára a kezébe került és érdekesnek találta. Gergő Sziveri Jánosról írt, nekem nemzedéktársamról és barátomról, a néhány évvel korábban rákos betegségben elhunyt újdéki költőről, akit akkor még nem sokan ismertek és olvastak Magyarországon. Ő igen.

Azóta eltelt több mint két évtized, hosszú és most mégis rettentően rövidnek bizonyuló idő, amit kiszabott még számára a sors, amit egymás közelében tölthettünk el, úgy, hogy tudtuk, bármikor összefuthatunk az egyetem folyosóján, vagy az épület előtt, én jövök kifelé, ő éppen köti le a kerékpárját, vagy fordítva, fején még ott a sisak, beszélünk pár szót, számtalanszor ismétlődő jelenet volt ez. És a megbeszélte találkozások, legelőbb a szemináriumok, azután a csütörtök esték, a szegedi klasszikus magyar irodalmi képzés emblematikus csütörtök estéi, műhelyszeminárium, ahol először hallgatóként bukkant fel, aztán doktorandusz lett, végül tanárként volt jelen. Vétésforgók, Kaposvártól Kolozsvárig, hosszú közös autótutak, útmenti fogadók, hosszú beszélgetések.

Ennek most vége.

Pedig együtt kezdtük volna ezt a szemesztert is, de szeptember 14-én, csütörtökön délelőtt a következő rövid üzenet érkezett tőle:

...tegnap váratlanul bent tartottak Deszken, mert légmellet találtak nálam. Ezért ma biztos nem tudok menni este. Barátsággal. Gergő

Azután még egy rövid levél jött tőle péntek délelőtt, javaslat a félév programjához, és én nem gondoltam volna, hogy ez lesz a búcsúlevél.

Nem volt hivalkodó ember, nem hivalkodott sem a tehetségével, sem a tudásával. Pedig mindkettőből volt neki bőven. Szerény volt, csendesen dolgozott, és nem személyével, hanem munkája eredményével, publikációival keltett feltűnést, lett Martinkó- és Kiss József-díjas, lett a klasszikus magyar irodalom jól ismert, sokat hivatkozott és megbecsült kutatója. Doktori disszertációja, mely később könyvformában *A magyar episztola a felvilágosodás korában. Műfaj- és médiatörténeti értelmezés* címmel jelent meg 2008-ban, alapműnek bizonyult. Elméleti igényesség és felkészültség társul benne a filológiai aprómunkával, a régi szövegekkel való szeretetteljes bíbelődéssel, a forrásfeltárás örömeivel. Ányos és társai, kolostorok lakói, és a délceg katona, Barcsay

Ábrahám. Nevezhetnénk ezt a szoros olvasás örömének is, megelégedhetett volna ezzel, a feltárt anyaggal még sokáig elbíbelődhetett volna. Hiszen vannak egytémás, kitűnő irodalomtörténészek is, nem is kevesen.

Labádi Gergely azonban egy idő után már nem a szoros olvasás örömét, hanem a korlátait érezte, és igyekezett áthágni azokat. Ez a törekvése egészen új utakra vezette. Az ő nemzedéke volt talán az első, amely már nem csupán intelligens írógépként nyúlt a számítógéphez, hanem az érdekelte, hogy ez az új technika milyen új észlelési módokra ad lehetőséget. Ehhez persze otthonosan kellett mozognia ebben a szép új világban, aminek mi, akik nagyon idegenül mozogtunk benne, legelőbb a praktikus előnyeit éreztük. Mert hamar kiderült, hogy ha valamilyen problémánk támad a gépünkkel, vírust kell irtani vagy egy új programot telepíteni, azt használni, akkor bátran lehet támaszkodni őrá. Nagyon empatikus és segítőkész volt mindig, olyan ember, akiből üdítően hiányzott a számítógépes szakemberek felsőbbrendűsége, melyet szeretnek éreztetni a számítógépet használó pórnépekkel szemben.

Később azonban kiderült, hogy jóval többről van szó ennél.

A legújabb angolszász és német szakirodalmat bűvárolva ő döböntett rá minket arra, hogy a digitális fordulat utáni világrend a szövegek megközelítésének radikálisan új formáit, egyelőre inkább a lehetőségeit is magával hozta. Mert a digitalizált és a világhálóra kötött számítógép előtt ülve kényelmesen lehívható szövegek nagy tömege nem csupán a régóta óhajtott világkönyvtár megszületését hozza látható közelségbe, hanem a szövegek értelmezésének új, a korábbiaktól alapvetően eltérő lehetőségét is. Mert ilyen mennyiségű szöveget ugyan mi, szoros olvasók, sohasem tudnánk egyben látni, de – nagyon durván leegyszerűsítve az általa nagyon cizelláltan körüljárt problémát – mi történik akkor, ha megtanítjuk a számítógépet, hogy olvassa el helyettünk a szövegeket és válaszolja meg a *mi* kérdéseinket?

Ez a felvetés az irodalomtörténeti kutatások módszertanának radikális újragondolását takarja, és Labádi Gergőnek köszönhetően a szegedi magyar irodalom tanszék nem utána kullogott, hanem együtt haladt azokkal, akik itthon és nagyvilágban a digitális kor új kérdéseire kerestek választ. *Digitális bölcsészet*, először furcsán hangzott, mára – neki köszönhetően is – bevett terminussá vált. Alighanem ő hívta fel elsőként a magyar kutatók figyelmét Franco Moretti nagyhatású fogalmára, a *distant reading*-re, majd emlegette azt együtt a Stephen Ramsay által bevezetett algoritmikus irodalomtudománnyal, és a Matthew L. Jockers által javasolt makroelemzéssel. „A lényeg persze – fogalmazott Jockers könyvének tanulmány értékű ismertetése során – mindhárom fogalom esetében ugyanaz: nagy mennyiségű adat rendszeres, következetes vizsgálata, amely révén olyan trendek, mintázatok azonosíthatók, amelyek a szövegek mikroszintű vizsgálata, azaz a szoros olvasás révén nem elérhetők.” A terjedelmes recenziót pedig azzal fejezte be, hogy akiket a technikai részletek is érdekelnek, elővehetik Jockers másikat, az ismertetett könyvnél egy évvel később megjelent munkáját.

Igen, ha ez érdekel bennünket, most már valóban a könyvre és magunkra vagyunk utalva.

Gergőt már nem kérheti fel az Irodalomtörténet szerkesztősége, hogy írjon róla is újabb ismertetőt, de nem állíthatjuk meg mi sem a folyosón, hogy megkérjük: a legközelebbi csütörtök estén beszéljen erről nekünk.

Pedig sokáig nem adtuk fel a reményt, hogy ezt megtehetjük, és nem adta fel a küzdelmet ő sem. Jól emlékszem arra a csütörtök estére – mindenki elment már, ketten maradtunk a sarokszobában, 2015 novembere volt – amikor elmondta nekem: foltot találtak a tüdején. Én, talán ösztönösen, a kisebbik rosszat igyekeztem érteni: azt hittem, tbc-ről van szó. Rám nézett, és azt mondta: nem.

Az a majdnem két év, amely ezután következett, heroikus volt. Mert igyekezett úgy élni, mintha semmi sem történt volna. Bejárt az egyetemre, egyre soványabban, a betegségtől és a nehezen megszerzett gyógyszerektől, ezektől a szörnyű mérgektől egyaránt elgyötörve. Egyre nehezebben járt, kínozták a köhögési rohamok. Mégis jött, sapkában és fehér cérnakesztyűben. Ha kellett, hihetetlen önfegyellemmel fojtotta vissza a köhögési ingert, és megtartotta az óráit. Azután bement a szobája, és köhögni kezdett, erősen, hosszan és félelmetesen.

Elképesztő módon tartotta magát egyébként is. Elkecseregettnak egyszer hallottam a hangját, 2015 karácsonya és újéve között, békés, lassú napok a két ünnep között, amikor azzal hívott fel, hogy daganatot találtak a koponyájában és rapid műtetre vár a klinika folyosóján. „Na, a docens!” – mondta, és elkecseregetésébe szarkazmus vegyült. De amikor a műtét után meglátogattam a klinikán, az elkecseregetésnek már nyomát sem találtam. Kijött velem a folyosóra, turbán a fején, beszélgettünk.

Élni akarása életben tartotta.

Velünk volt csütörtök esténként, tavasszal eljött velünk Pécsre. Sokak számára emlékezetes, szép előadást tartott a budapesti Berzsenyi-konferencián. Augusztusban pedig, hogyléte iránt érdeklődő levelemre így válaszolt:

Szia,

most jöttünk vissza egy párnapos gyulai nyaralásról, ahol megszagolhattam egy parfümöt az Almásy-kastélyban, az 1846-ban megalkotott „nemzeti illat”-ot, elég kellemes volt. Egyébként pedig tanulmányokat írok, amelyekben a gépi olvasás eredményeit értelmezem.

Amióta tudósítottalak, nem volt újabb vizsgálat, holnap döntik el, hogyan menjen tovább a kezelés...

Barátsággal. Gergő

Újabb leveleiben a *barátsággal* szó után mindig pontot tett.

Tavasza óta a vizsgálatok eredményei rendre kedvezőbbek voltak, és mi egyre inkább reménykedtünk. Aztán jött az utolsó hétvége, és hirtelen minden rosszra fordult, és két nappal később mindennek megfoghatatlanul vége lett.

Hosszú felkészülési időt küzdött ki nekünk, de nem voltunk felkészülve mégsem.

Nincsen több szavam.

Gergő, drága barátom, legyen neked könnyű a föld.

Szajbély Mihály

A temetésen elhangzott búcsúbeszéd megjelent a *Helikon* 28. kötetének 20. számában (2017).