

Pázmány Péter Katolikus Egyetem
Bölcsészet- és Társadalomtudományi Kar

**A digitális filológia Magyarországon:
elvek és gyakorlatok**

Doktori (PhD) értekezés tézisei
Fellegi Zsófia

Irodalomtudományi Doktori Iskola

Témavezető: Dr. Almási Zsolt tanszékvezető egyetemi docens

A doktori iskola vezetője: Dr. Dobos István tanszékvezető egyetemi tanár

Budapest

2024

1. A kutatás előzményei, problémafelvetés

A digitális médiumváltás hatással van a jelen kor minden kulturális gyakorlatára, a tudományra, azon belül a bölcsészetre, az irodalomtudományra, a filológiára is. A nemzeti kulturális örökség filológiai szempontból legmagasabb színvonalú szövegkiadásai a kritikai kiadások, a filológia, azáltal, hogy ellenőrzött, „stabil” szövegkorpuszokat állított elő a bölcsészeti kutatások számára, mindig is formálta azt, hogy mi válik a kutatás tárgyává. Az elmúlt évtizedekben a számítástechnika és az internet fejlődése magával hozta a tudományos módszertanok változását is. Az irodalomtudomány és a filológia területén is felmerült az igény olyan kutatási infrastruktúra kifejlesztésére, amely lehetővé teszi a fizikai határokon átívelő kutatási hálózatok kiépítését.

Jelen disszertáció nem vállalkozik a magyar digitális filológia történetének megírására, ez a vállalkozás önmagában egy nagymonográfia témája lehetne. A kutatás során az is egyértelművé vált, hogy a magyarországi digitális filológiai műhelyek és projektek genezise csupán töredékesen dokumentált, számos kulcsfontosságú információ nem áll rendelkezésre, az ehhez szükséges feltáró munkálatok meghaladják a disszertáció kereteit. A disszertáció szintén nem nyújtja a digitális filológia rendszeres elméleti áttekintését. Ez a feladat talán még az előzőnél is nagyobb, már csak a technológia döbbenetesen gyors változása miatt is lehetetlen lenne monográfiászerűen egy ilyen problémát tárgyalni. A disszertáció ugyanakkor érinti a magyar digitális filológia történetének egyes fejezeteit és elméleti szempontokat is tárgyal, de csak egy konkrét digitális filológiai műhely, a DigiPhil (A magyar irodalomtudomány filológiai portálja) gyakorlatának kontextusában.¹ A disszertáció elsődleges célja az, hogy segítsen a magyarországi filológia területén felmerülő gyakorlati és elméleti problémák áttekintésében, a jó gyakorlatok kiemelésével. Ennek háttérében az a törekvés áll, hogy a DigiPhil körül az elmúlt egy évtizedben összegyűlt tapasztalatokat más filológiai műhelyek és, különösen, a jövő digitális filológusai számára közvetítse. A disszertációban tehát arra vállalkozom, hogy a DigiPhil genezisének és jelen gyakorlatának problémaorientált áttekintését nyújtsam. Ehhez, egyrészt, elengedhetetlen a DigiPhil-t megelőző, illetve a kialakulása körüli magyarországi és nemzetközi kontextus vázlatos bemutatása, valamint azoknak az elméleti kérdéseknek, vitáknak, kurrens diskurzusoknak a megidézése, amelyek közvetlenül hatással voltak és vannak a DigiPhil kutatási módszertanának mikro- és makrodöntéseire. A disszertáció egyaránt

¹ A DigiPhil szolgáltatás elérhetősége: <https://digiphil.hu/>.

tárgyalja a DigiPhil 1.0 és 2.0 gyakorlatát és azokat az elméleti kontextusokat, amelyek szerepet játszottak ezek kiépítésében.

2. A kutatómódszertan és a dolgozat szerkezete

A disszertáció bevezető fejezete két részre osztható. A fejezet első felében bemutatom azokat az alapfogalmakat, amelyekre az érvelés támaszkodik:

- digitális filológia (és digitális bölcsészet)
- digitális és elektronikus (ritkábban hálózati) kiadás
- informatizált kiadás
- born digital kiadás
- jelölőnyelvi átírat
- adatbázis, adatszerkezet

Bár a disszertáció célja nem a magyarországi digitális filológia történetének megírása volt, ahhoz, hogy a későbbi fejezetek, elsősorban a 2014-ben induló DigiPhil egyes kiadásainak példái köré épülő érvelés történeti kontextusba kerüljön szükség volt egy vázlatos és szelektív áttekintésre az érvelés szempontjából legjelentősebb digitális filológiai fejleményekről az elmúlt két évtizedből, a bevezető fejezet második része erre az áttekintésre összpontosít.

A disszertáció első fejezete (*Átmenet az analógból a digitális filológiába*) az analógból a digitális filológiába való átmenet metafilológiájával foglalkozik (elméleti szempontok figyelembevételével), de konkrét, a DigiPhil csoport szakmai támogatásával közzétett szövegkiadások összefüggésében. A bemutatott digitális tudományos és kritikai kiadások sokfélék, a fejezet arra törekszik, hogy a különböző filológiai iskolák, illetve szövegkorpuszok számbavételével és elemzésével áttekintést nyújtson a digitális filológia legfontosabb kérdéseiről. Elsőként nyomtatott kiadások informatizálására láthatunk példát, itt a digitalizálás és informatizálás különbségére kerül a hangsúly. Ebben a részben az Arany János kritikai kiadások és Madách Imre *Az ember tragédiája*, *Szinoptikus kritikai kiadás* retrokonverzióval történő informatizálásának bemutatásán keresztül kaphatunk képet az egyes kiadástípusok eltérő szövegközlési elveinek egy rendszerben történő leképezéséről. Ezt követi a *born digital* kiadások elemzése. A *born digital* kiadások esetén nem tekinthetünk el a digitális filológia a szövegkiadás elveire gyakorolt hatásának, valamint a filológusi szerep megváltozásának

elemzésétől. Az elemzés több kiadástípust vizsgál, az első esetben (*Móricz Zsigmond levelezésének kritikai kiadása*) a *born digital* kiadás nem egy sorozat részeként készült, a szövegközlési elveknek nem kellett egy sorozat elveihez igazodni. A második esetben (*Oláh Miklós levelezésének kritikai kiadása*, 2. kötet) a *born digital* kiadás előzménye, az első kötet, egy célszoftver segítségével készült, amelynek segítségével a sajtó alá rendező a kiadás készítésének teljes folyamatát, az átírástól kezdve az kritikai apparátus beillesztésén át a névmutató és a tördelt nyomdai formátum előállításáig egy rendszerben végezte el. Az első kiadás elkészítéséhez használt célszoftverben lehetőség lett volna szabványos jelölőnyelvi átírat készítésére, azonban digitális filológiai kompetenciák nélkül az ehhez szükséges beállítások nem végezhetőek el. A második kiadás egy, a kiadás igényeire szabott irodai szoftvert imitáló környezetben készült, ahol a felhasználóbarát felület segítségével jelölőnyelvi átírat jött létre, anélkül, hogy a sajtó alá rendezőnek ismernie kellett volna a TEI szabványt és az XML nyelvet. A kiadás a *Bibliotheca Scriptorum Medii Recentisque Aevorum* sorozat része, így a jelölőnyelvi átírat kialakításakor mind a sorozat, mind a DigiPhil szövegközlési elveit figyelembe kellett venni. Az utolsóként bemutatott vállalkozás, *Babits Mihály Összes Művei* sorozat Magyarországon biztosan, de valószínűleg európai viszonylatban is egyedülállónak számít. A genetikus kritika elveit követő kiadássorozat most készülő köteteinek XML reprezentációja, az anyag természetéből fakadóan, rendkívül bonyolult, a legújabb, a szemantikus web szemléletmódjának elterjedésével létrejött technológiák alkalmazása nélkül lehetetlen vállalkozás lenne elkészíteni. A szemantikus web, azaz a Web 3.0 koncepciójának alapját az a jelenség adta, hogy a Web 1.0 és 2.0 elterjedésével már a 2000-es évek elején is kezelhetetlen mennyiségű, rendezetlen adat és dokumentum vált elérhetővé a világhálón. A koncepció szerint a problémát intelligens gépi megoldásokkal lehet megoldani, azonban ehhez az adatoknak gépek által is értelmezhető formátumban kell rendelkezésre állniuk.

A Digitális filológia és a szemantikus web című fejezet ezen technológiák és a szemantikus web szemléletmódjának a digitális filológiára gyakorolt hatását vizsgálja. Ez a hatás egyaránt tetten érhető a digitális filológia elméletében és a gyakorlatában is. A disszertáció a nemzetközi kísérletek és új elméletek elemzésén keresztül mutatja be a szemantikus web elvét és működését. A DigiPhil projekt a digitális kiadásokat a kezdetektől bibliográfiai és biográfiai adatokat tartalmazó háttéradatbázissal kötötte össze, az adatbázis könyvtári rendszerben épült, amely számos kompromisszummal járt. A szemantikusweb-technológiának alkalmazásával olyan gráf formájú tudástár hozható létre, amely már nem csak a könyvtári rendszerben tárolt

adatokat képes kezelni, hanem képes a kutatási adatokat szemantikus hálózatba rendezni és az adatokból további következtetéseket levonni.

A digitális kiadásokon túl: Jelölőnyelvi átiratok elemzése című fejezet azt vizsgálja, hogy a kritikai kiadások XML reprezentációi milyen számítógépes eszközökkel elemezhetők. A digitálisan reprezentált irodalmi szövegek esetében a kutatás sikerességéhez szükség van a metaadatok és a szövegek együttes kezelésére, illetve a szöveg több szempontú gépi feldolgozására. A nemzetközi szabványok és ontológiák, névterek használata révén lehetőség nyílt az egymástól – fizikailag – távol eső adatbázisokban való értelmes és kifinomult közös keresésre. A fejezet első része, a kurrens web teóriák figyelembevételével, a metaadat-aggregáció folyamatát mutatja be. A fejezet második része jelölőnyelvi átiratok hálózati vizualizációinak lehetőségeivel foglalkozik. Az elemzések elvégzésére különböző eszközök állnak rendelkezésre, a fejezet ezen részében az *Arany János Levelezése 1866–1882* informatizált kiadásának, azon belül is a levelek metaadatainak feldolgozását mutatom be, majd a metaadatok különböző eszközökkel történő vizualizációs lehetőségeit hasonlítom össze. A DigiPhil indulása óta közel 10.000 XML fájlból álló korpusz épült. A gyűjtemény jelentős részét 19–20. századi szerzők munkái teszik ki. A korpusz méretéből és a szerzők időbeli közelségéből adódik a lehetőség, hogy a rendelkezésre álló kritikai kiadásokon nyelvstatisztikai elemzéseket végezzünk. A szövegstatisztikai módszerekkel olyan, adott szerzőre jellemző mintázatokat deríthetünk fel, melyek korábban sokszor láthatatlanok maradtak a kutatók számára. A fejezet végén Kosztolányi Dezső kritikai kiadásainak vizsgálatával mutatom be a rendelkezésre álló módszerek némelyikét.

A digitális filológia térnyerése az elmúlt évtizedekben vitathatatlan, azonban hiába az évtizedes tapasztalatok, a digitális filológia oktatása kiforratlan még, nem áll rendelkezésre átfogó képzés a leendő szakértők számára. Kifejezetten időszerű kérdés, tekintve, hogy a digitális, illetve digitalizált kiadásokkal foglalkozó kutatócsoportok száma egyre növekszik, de a magyarországi képzések jelenleg nem gondoskodnak szakértői „utánpótlásról”. A disszertáció záró fejezete reflektál erre a problémára és bemutatja, milyen lehetőségek álltak rendelkezésre az elmúlt években és manapság a digitális filológia oktatása terén, illetve milyen szempontokat érdemes figyelembe venni egy komplex képzés kialakítása során. A javasolt szempontok egyaránt érvényesíthetők a graduális és a posztgraduális képzésben.

3. Eredmények, konklúzió

A disszertáció egyik fő eredménye a digitális filológia területén alkalmazott terminológia körül határolása, a szakdiszciplína területén használt fogalmak definiálása. A disszertációban arra vállalkoztam, hogy a DigiPhil genezisének és jelen gyakorlatának problémaorientált áttekintését nyújtom, ennek megvalósítása során sikerült a korábban elszórtan megtalálható digitális filológiai tudást és tapasztalatokat rendszerezni. A disszertáció áttekintést nyújt a DigiPhil döntéseinek háttérben álló elméleti megfontolásokról, ezen felül távlati képet nyújt a digitális filológia határterületeiről, a jelölőnyelvek számítógépes elemzésének lehetőségeiről. A disszertáció reflektál a digitális filológia területén felmerülő elméleti és módszertani kérdésekre. Magyarországon jelenleg nem érhető el átfogó digitális filológia oktatás, a dolgozat, számba véve a korábbi jó gyakorlatokat, javaslatokat fogalmaz meg rövidebb és hosszabb távú képzések kidolgozására.

4. A témában végzett publikációs tevékenység

DOBÁS Kata, FELLEGI Zsófia, PALKÓ Gábor. „A kis gömböc meséje – az ITIdata irodalomtudományos adatbázis fejlesztése 2022–2023-ban”. In *Új technológiákkal, új tartalmakkal a jövő digitális transzformációja felé: 32. Networkshop: országos konferencia: 2023. április 12-14. Pannon Egyetem, Veszprém*, szerkesztette TICK József, KOKAS Károly és HOLL András, 192–198. Budapest: Hungarnet Egyesület, 2023.

<https://doi.org/10.31915/NWS.2023.28>

FELLEGI Zsófia. „WIKIDATA – ELTEdata – WIKIMEDIA: A szövegkiadások technikai háttere”. In *A tudományos tudásáramlás mintázatai a Magyar Királyságban, 1770-1830: DigiTudásáramlás (DITU) /DigiCirculation of Knowledge (DICIKO)*, szerkesztette BALOGH Piroska, BODNÁR-KIRÁLY Tibor, GURKA Dezső, KRÁSZ Lilla, KURUCZ György, SZ. KRISTÓF Ildikó, Budapest: ELTE BTK Digitális Bölcsészeti Tanszék, 2022.

<https://doi.org/10.5281/zenodo.7584606>

SZEMES Botond, BAJZÁT Tímea, FELLEGI Zsófia, KUNDRÁTH Péter, HORVÁTH Péter, INDIG Balázs, DIÓSSY Anna, és mtsai. „Az ELTE Drámakorpuszának létrehozása és lehetőségei”. In *Valós térben - Az online térért: Networkshop 31: országos konferencia. 2022. április 20–22. Debreceni Egyetem*, szerkesztette TICK József, KOKAS Károly és HOLL András, 170–178. Budapest: Hungarnet Egyesület, 2022. <https://doi.org/10.31915/NWS.2022.21>.

FELLEGI Zsófia. „A digitális filológia infrastruktúrái. A DigiPhil megújulásáról”. In *Valós térben - Az online térért: Networkshop 31: országos konferencia. 2022. április 20–22. Debreceni Egyetem*, szerkesztette TICK József, KOKAS Károly és HOLL András, 338–344. Budapest: Hungarnet Egyesület, 2022. <https://doi.org/10.31915/NWS.2022.43>.

FELLEGI Zsófia. „Átmenet az analógból a digitális filológiába: médiumváltás? Madách Imre: Az ember tragédiája. Szinoptikus kritikai kiadás”. *Helikon* 67, 1. sz. (2021): 84–102.

FELLEGI Zsófia. „A digitális filológia oktatása: Gyakorlati tapasztalatok”. In *Digitális eszközök a középiskolai irodalomoktatásban*, szerkesztette MOLNÁR Gábor Tamás és FEJES Richárd, 157–178. Budapest: Prae.hu, 2021.

FELLEGI Zsófia és PALKÓ Gábor. „Arany-kéziratok és kritikai kiadások közzététele az Arany János Emlékévben”. *Helikon*, 66. sz. (2020): 82–98.

FELLEGI Zsófia. „Digital Philology on the Semantic Web. Publishing Hungarian Avant-Garde Magazines”. In *The (Web)Sites of Memory*, edited by Donald E. MORSE, Zsófia RÉTI és Miklós TAKÁCS, 105–116. *Loci Memoriae Hungaricae* ; 8. Debrecen: Debrecen University Press, 2018.

FELLEGI Zsófia és VADÁSZ Noémi. „Digitálisbölcsészet-oktatás Magyarországon”. In *Digitális (szöveg)kultúrák a bölcsészképzésben*, szerkesztette L. VARGA Péter, MOLNÁR Gábor Tamás és PALKÓ Gábor, 168–180. Budapest: ELTE-BTK, 2018.

CSÉVE Anna, FELLEGI Zsófia és KÓMÁR Éva. „Móricz Zsigmond levelezésének (1892–1913) digitális kritikai kiadása”. *Digitális Bölcsészet* 1 (2018. szeptember 19.): 159–174. <https://doi.org/10.31400/dh-hun.2018.1.227>.