

Digitális szövegkiadások és kutatástámogatás a bölcsészettudományokban: A dHUpla platform bemutatása

The dHUpla Platform: Digital Text Editions and Research Support in the Humanities

Varga Emese
varga.emese@oszk.hu
digitálistartalom-fejlesztő, OSZK
Digitális Bölcsészeti Központ

Beérkezett: 2024.03.18.

Elfogadva: 2024.04.23.

Publikálva: 2024.07.11.

Az Országos Széchényi Könyvtár Digitális Bölcsészeti Központjának digitális filológiai platformja, a dHUpla egyik fő célja a közgyűjteményekben őrzött, főként magyar nyelvű kéziratos források digitális feldolgozása és közzététele. A szolgáltatás egységes, szabadon hozzáférhető kutatói környezetet kíván teremteni az irodalomtudomány, a nyelvtudomány és további humán tudományterületek kutatói számára.

Jelen tanulmány a platformon publikált szövegkiadásokon keresztül a digitális szövegkiadások kutatástámogató szerepét állítja fókuszba, kiemelve a médiumváltás adta új szövegkezelési lehetőségeket, illetve a kiadások mögött zajló munkafolyamatok előnyeit és kihívásait.

digitális filológia, digitális szövegkiadások, elektronikus szövegkiadások, kutatástámogatás

The Digital Humanities Platform, also called dHUpla, is developed at the Digital Humanities Centre of the National Széchényi Library, and aims to process and publish manuscript resources, mainly in Hungarian, held in GLAM sector institutions. The service aims to create an unified, freely accessible research environment for researchers in literary studies, linguistics and other humanities.

This paper focuses on the research support role of digital text editions, highlighting the new text management opportunities offered by the change of the medium, while referring to the benefits and the challenges of the text editions workflow.

digital philology, digital text editions, electronic text editions, research support

1. A dHUpla platform bemutatása

Az Országos Széchényi Könyvtár Digitális Bölcsészeti Központjának (továbbiakban DBK) digitális filológiai platformja, a dHUpla¹ (<https://dhupla.hu/>) egyik fő célja a közgyűjteményekben őrzött kéziratos források digitális feldolgozása és közzététele. A szolgáltatás egységes, szabadon hozzáférhető kutatói környezetet kíván teremteni az irodalomtudomány, a nyelvtudomány és további humán tudományterületek kutatói számára. A platformon közzétett anyagok egyrészt népszerűsítő jelleggel a nagyközönséget célozzák meg, ugyanakkor az irodalomtudományi, történeti, társadalomtudományi kutatások számára kézenfekvő

Cite as/Így hivatkozd: Varga, E. (2024) Digitális szövegkiadások és kutatástámogatás a bölcsészettudományokban: A dHUpla platform bemutatása / The dHUpla Platform: Digital Text Editions and Research Support in the Humanities, Central European Library and Information Science Review (CELISR), 1(2), p. 134–143.
<https://doi.org/10.3311/celistr.37709>

forrásanyagként szolgálnak, alkalmasak számítógépes szövegelemzésekre és az előállított adatok alapján különböző szempontú adatvizualizációk készítésére is.

A dHUpa felülete három tematikus aloldalból áll:

- szövegkiadások menüpont alatt érhetők el a digitális publikációk,
- az entitástár almenü a kiadásokhoz kapcsolódó névelemeket tartalmazza (személynevek, műcímek),
- a kreatív oldalon szereplő tartalmak a digitális szöveg létrehozása során előállított adatokat dolgozza fel, amelyek a kiadások készítésével párhuzamos kutatómunka eredményeként jöttek létre (pl. kapcsolati gráfok, stilometriai elemzések, szöfelmők).



1. ábra: dHUpa nyitó oldal

Minden a dHUpán megjelenő szövegkiadásnak egységes a struktúrája. A kiadásokhoz egy-egy gyűjteményleírás tartozik, amelyben a feldolgozott kéziratanyagról és az átírás elveiről olvashatunk, ez segíti a felhasználót a gyűjteményben való tájékozódásban is. Egy gyűjteményt a kiadás szempontjából összetartozó olyan digitális objektumok összessége jelenti, amelyek listában jelennek meg a felületen. Az objektumokat (ez lehet egy levél, egy naplófejezet, egy folyóirat-szám) több szempont szerint rendezhetjük, kereshetünk a teljes gyűjteményben, és kombinált szűrőket (ún. facetek) is alkalmazhatunk. Egy dokumentum megnyitáskor a gyűjteményi rekordok metaadatai külön ablakba kerülnek. Az átíratok a faksimile képpel együtt tanulmányozhatók, különböző módon:

- az osztott képernyős nézetben párhuzamosan láthatjuk a faksimilét és az átírt szöveget,
- alapértelmezett módon az átírat fölött jelennek meg a faksimile kép zónái – soronként és nagyobb kontextusban is. Ez az ún. text-image linking technológiának köszönhető, amely a Transkribus² szoftverben szegmentált és átírt digitális objektumok esetében valósul meg. A platform célzottan felhasználóbarát felépítésű, hogy a különböző funkciók segítsék az olvasót a gyűjteményben való eligazodásban, illetve a szöveg és a kép különböző megjelenítési módjai illeszkedjenek a kutató aktuális céljaihoz.

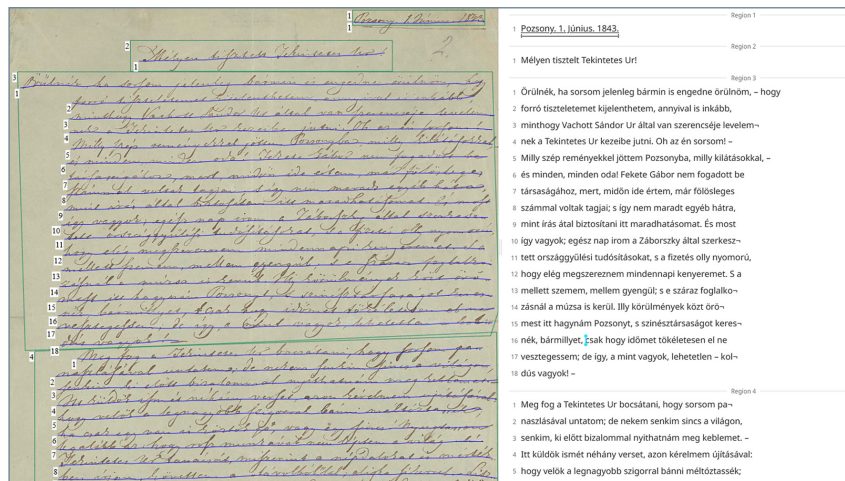
2. A digitális faksimilétől a digitális szövegig

Jelenlegi munkafolyamatunk első lépéseként a digitális faksimilét feltöltjük a Transkribus szoftverbe, amely egy felhasználóbarát szerkesztői felületet nyújt a kéziratok képi szegmentálására, átírására és annotálására. Az eszköz mesterségesintelligencia-alapú és saját kézírásfelismerő modell építésére is alkalmas, ezzel jelentősen meggyorsítva az átírás folyamatát. Ez a DBK által fejlesztett első magyar

"A kiadásokhoz egy-egy gyűjteményleírás tartozik, amelyben a feldolgozott kéziratanyagról és az átírás elveiről olvashatunk..."

"A TEI nemzetközileg támogatott standardja keretet biztosít a szövegek XML-szerkesztéséhez, hogy azok elemezhető, kereshető és megjeleníthető legyenek a digitális kiadásokban."

kézírásfelismerető modell 2021-ben vált publikussá (Szűcs 2021, p. 73–80). Az átiratokat a TEI³ nemzetközi szabványának megfelelő XML-formátumban exportáljuk a Transkribus szoftverből, majd metaadatokkal és szövegközi annotációkkal tovább gazdagítjuk az Oxygen XML Editor⁴ programban. A TEI nemzetközileg támogatott standardja keretet biztosít a szövegek XML-szerkesztéséhez, hogy azok elemezhető, kereshető és megjeleníthető legyenek a digitális kiadásokban.



2. ábra: minta a kézírásfelismerő Transkribus szoftver működésére (részlet Petőfi Sándor Bajza Józsefhez írt leveléből)⁵

A TEI XML egy széles körben használatos jelölőnyelv. A hosszú távú megőrzés szempontjából fontos, hogy az információkat olyan formátumokban tároljuk, amelyek használata mára elterjedt, mert ez növeli annak esélyét, hogy a formátum elavulásakor megfelelő konverziós programok álljanak rendelkezésre (Bia 2023, p. 291). A TEI szabványának megfelelő kódolással, az adott szabályrendszer követve, egy helyen őrizhető a szöveg és a hozzá tartozó metaadat-leírás, így külső dokumentációra való hivatkozás nélkül tárolhatjuk az információkat. (Bia 2023, p. 291–292). Számos előnye miatt vált napjainkra a TEI XML a digitális szövegkiadások szabványává (Stäcker 2020, p. 2–3).

Hazai viszonylatban az elmúlt évtized kiadásai már zömmel a TEI XML szabványa szerint készültek (Maróthy 2020). Palkó Gábor szerint „a digitális filológiában évtizedek alatt a TEI XML egyeduralgódóvá vált” (Palkó 2016, p. 1320). Így napjainkra az általánosan használt jelölőnyelvként is hivatkoznak rá (Debreczeni 2020, p. 405). Mindezek ellenére a TEI standardját több kritika is éri. Egyik gyakori probléma, hogy az XML-ben körülményes az átfedések jelölése, ami nehézséget okoz az összetett (például nyelvi és szemantikai) jelenségek kódolásában és megfelelő megjelenítésében (Stäcker 2020, p. 2).

A másik hátrány a keretrendszer rugalmassága, amely egyes vélemények szerint már a standard helyes működését is veszélyezteti. Debreczeni szerint „továbbra is vita folyik abban a tekintetben, hogy milyen mértékű lehet és legyen a TEI–XML szabványosítása. A TEI rugalmassága főként abban áll, hogy ugyanazt a jelenséget többféleképp lehet kódolni, és ezért nehéz meghatározni, hogy adott esetben melyik legyen az általános eljárás.” Debreczeni rávilágít arra, hogy ez a rugalmasság jelenti a TEI–XML standard lényegét: „Ahogy a nyomtatott kritikai kiadások számára nem írhatunk elő részletekbe menően egységes és kötelező tartalmi formátumot, úgy szükséges a jelölőnyelvi struktúra részleteit is az anyag sajátosságaihoz szabni, megmaradván persze a TEI–XML standard keretei között, annak belső logikájából építkezve, egy adott kiadásban belül következetesen” (Debreczeni 2020, p. 405–406).

A TEI rugalmasságának problémájával a dHUpa honlapján megjelenő szövegek esetében is szembe kellett néznünk. „Az általános elvekhez képest meglehetősen szigorú szabályrendszer létrehozására volt szükség ahhoz, hogy az annotációk alapján a platformon való megjelenítés, kereshetőség, lekérdezhetőség stb. jól működjön” (Mihály 2022, p. 348). Ehhez a DBK elkészítette a TEI-XML specifikációt, amely meghatározza azokat a címkéket és címkézési szabályokat, amelyek az egyes szövegtípusoknál előfordulhatnak.⁶ Bár ez egy kötöttebb, szigorúbb szabályrendszert épít, a szövegjelölési opcióknak továbbra is széles eszköztára áll rendelkezésre, amely a feldolgozandó forrásszöveg sajátosságaira szabható.

Az Oxygen ilyen szempontból rugalmasabb, lehetőséget biztosít ún. framework fejlesztésre, amely egy személyre szabott szerkesztői felületet nyújt, eszköztárakkal, menüvel, specializált megjelenítéssel és ellenőrzési lehetőségekkel. A framework a DBK által készített specifikáción alapszik, és egy olyan kezelési felületet nyújt, amelyet bárki könnyen használhat a TEI XML struktúra és címkékészlet mélyebb ismerete nélkül is. Ez biztosítja, hogy a dHUpára szánt szövegkiadásokat bárki hatékonyan tudja szerkeszteni, megteremtve a lehetőségét annak, hogy az infrastruktúra külső kutatók és intézmények projektjeit is befogadja. A szövegközi jelölések mellett az Oxygen frameworkben történik a TEI-fejléc (Header) kitöltése is, ahol a digitális objektumok metaadatait rögzítjük. Az adatgazdagítást a keretrendszer olyan automatizált műveletei segítik, mint a mesterségesintelligencia-alapú névelem-felismerés és a beépített névelem-azonosítás (vagyis a névelemek külső névterekkel való összekötése). Az infrastruktúra középpontjában a GitLab⁷ verziókövető rendszere áll, amelyben a tárolt anyagok valamennyi változata elérhető és visszakövethető, szintén a GitLabból indul ki minden automatizált folyamat, és ezzel áll kapcsolatban közvetlenül vagy közvetve az összes használt eszköz (Mihály 2022, p. 347).

A szerkesztés alatt álló szövegkiadások megjelenésének és működésének ellenőrzését egy belső felhasználású, ún. staging-szerver biztosítja, ahol nemcsak a publikált, hanem minden repozitóriumban található (pl. feltöltés alatt álló „draft”) tartalom is megtekinthető. Így a szövegkiadásban részt vevő munkatársak már a tartalom-előállítás folyamata során is láthatják eredményeiket, a javítandó hibákat, illetve jelezhetnek informatikai teendőket is (Mihály 2022, p. 352). A szerkesztő a befejezett fájlok státuszát a TEI jelölésben módosítja, ami által automatikusan lefutnak a publikáláshoz szükséges transzformációk, és a fájlok megjelennek a nyilvános felületen.

3. A különböző szintű és típusú digitális szövegkiadások szerepe és funkciója

A digitális vagy elektronikus szövegkiadásoknak egyre szélesebb a szakirodalma hazánkban is. Maróthy Szilvia *Digitális bölcsészet – a szövegtől az adatig* című kötetében a fogalom különböző meghatározásait járja körbe a magyar és az angolszász szakirodalomban. (Maróthy 2020). Megállapítása szerint, miképp az MTA Textológiai Munkabizottság által 2004-ben közreadott kiadási Alapelvek négyféle szövegtípust különböztet meg – a kritikai kiadást, a forráskiadást, a genetikus kiadást és a népszerű kiadást –, ugyanúgy beszélhetünk elektronikus kritikai kiadásról (főszöveggel), elektronikus forráskiadásról, elektronikus genetikus kiadásról és elektronikus népszerű kiadásról is (Maróthy 2020).

A dHUpán megjelenő szövegkiadások a felsoroltak közül három kiadástípusra 1. a digitális forráskiadásra, 2. a digitális kritikai kiadásra és 3. a digitális népszerű kiadásra adnak példát, eltérő tudományos szintű, illetve célú publikációk formájában, ezért átírási és textológiai jelölési elveik is különböznek egymástól.

"Az adatgazdagítást a keretrendszer olyan automatizált műveletei segítik, mint a mesterségesintelligencia-alapú névelem-felismerés és a beépített névelem-azonosítás..."

1. A digitális forráskiadás eddig publikálatlan kéziratos források közzétételét jelenti, amely betűhű szövegközlésre törekszik és alapvető textológiai annotációkkal rendelkezik (pl. törlés, rövidítés, betoldás, hiba), részletesen metaadatolt, illetve a kiadás céljainak megfelelően adatgazdagított (pl. tulajdonnevek beazonosítása). A metaadatstruktúrát úgy alakítottuk ki, hogy az adott szöveg típusra vonatkozó minden lehetséges adat rögzíthető legyen a TEI Headerben. Vannak kötelezően kitöltendő és opcionális elemek, amelyek az adott kiadás igényeire szabhatók. Így egy levelezés esetén a keletkezés, a küldés, az átvétel adatai mellett a kézirat részletes leírására, a szöveg típusára, a korábbi megjelenések és kritikai jegyzetek rögzítésére is van lehetőség, emellett a helynevek és személynevek közvetlenül a frameworkből külső névterekkel köthetők össze. Példaként említeném Kiss József levelezésének digitális forráskiadását,⁸ amely első lépésben a Petőfi Irodalmi Múzeumban (PIM) található kéziranyag feldolgozásán és közzétételén alapszik. A források átírásakor a betűhű elvek követése érvényesült, kivételt csak a megértést hátráltató esetek képeznek. A szövegeket alapvető textológiai annotációkkal láttuk el (pl. törlés, rövidítés, betoldás, hiba), illetve a gyűjteményi rekordok metaadatait is feltüntettük. A Vershaza-gyűjtemény⁹ egyedi módon az OSZK Esszencia című állandó kiállításának Vershaza-termében található kéziranyagán alapszik, amely a Himnusz, a Szózat és a Nemzeti dal fennmaradt kéziranyagát, legfontosabb szövegkiadásait, a megzenésítéseket és a különböző fordításokat állítja fókuszba. A digitális forráskiadás alapvető célja, hogy a kiállítás kéziratos anyagát (a betűhű átírást a kéziratképekkel együtt) tanulmányozhatóvá tegye, továbbá a digitális szövegkiadás lehetőségeit kihasználva arra is törekszik, hogy a lehető legtöbb kapcsolódó szövegkiadást, online tartalmat, hanganyagot kössön össze a forrásszövegekkel.

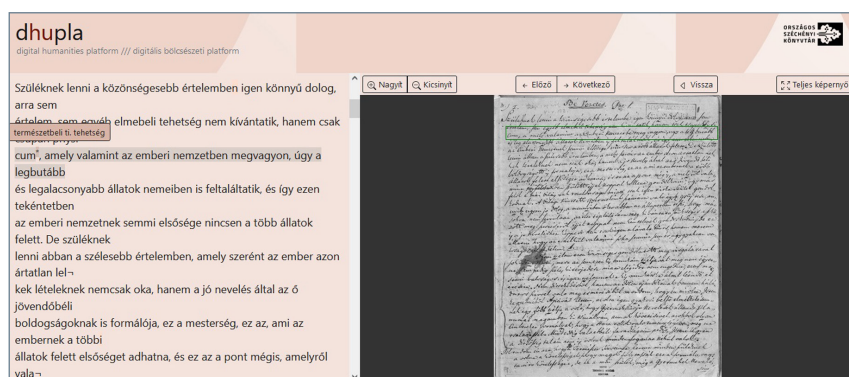


3. ábra: Vörösmarty Mihály: Szózat fogalmazvány (1. változat),
Forrás: dHupla Verhaza-gyűjtemény)¹⁰

2. Digitális kritikai kiadás formájában született Móricz Zsigmond levelezésének¹¹ forrásfeltárása és közzététele, amely az annotációk, az adatgazdagítás és szö-

vegkritikai jegyzetek tekintetében gazdagabb egységet képez. A szövegben említett entitások (személynevek, földrajzi nevek, intézménynevek, műcímek) beazonosítása, gyűjteményspecifikus jegyzetekkel való ellátása, illetve külső névterekkel összekötése a kutatói munka fontos részét képezi, mint ahogy a levélírókra, címzettekre és a datálásra vonatkozó kritikai jegyzetek írása is.¹²

3. A digitális népszerű vagy népszerűsítő kiadás nem törekszik a betűhű átírásra, az olvashatóság érdekében bizonyos szövegjellemzők modernizáltak, a mai helyesírás szabályait követik. Teleki László tanácsadása a nevelésről¹³ dHUplán közölt kézirata Fábri Anna tanulmánykötetének¹⁴ forrasszövegét képezi. A szövegátírás tekintetében fellelhető lényeges változtatások az eredeti kézirat-hoz képest, ezek nemcsak az aposztrófok elhagyásában, a kis- és nagybetűk, az egybe- és különírások mai szabályaihoz történő lehetőség szerinti igazodásban, a központosítás módosításában nyilvánulnak meg, hanem a fonetikus írásmód csaknem teljes elhagyásában is. Így a szövegközlés gördülékenyebbé teszi az olvasást, az idegen (többségében latin) eredetű szavakhoz fűzött jegyzetek, illetve egy tematikus „Teleki–latin szótár” tovább segítik a szövegértést.



4. ábra: Teleki László tanácsadása a nevelésről, minta a beépített idegen szavak szótára¹³

Minden kiadáshoz különböző szűrési szempontok állíthatók be a dHUpla felületén, amelyek az adott szövegtörzs XML-alapú, metanyelvi kódolásán alapulnak. Ezek egyaránt vonatkozhatnak a szövegben jelölt textológiai jelenségekre és entitásokra, illetve a TEI Headerben rögzített metaadatokra is. A Teleki-kiadásban például (a forrasszöveg természetéből adódóan is) kevesebb hangsúly került a szűrési szempontok beállítására, itt a forrasszöveg és az eredeti kézirat vizsgálatára került a fókusz. A Kiss József-levelezésben viszont többféle kombinált szűrőt alkalmaztunk: a levelek a típusuk, a levélírók, a címzettek, a keletkezés és feladás helye, valamint a keletkezés éve szerint is leválogathatók. A Móricz-levelezés-kiadáshoz jelentős háttérkutatás társul, ami a szűrési funkciók gazdagságában is megnyilvánul. A Kiss József-kiadásban említett szűrési feltételek mellett a Móricz-levelek az említett műcímek és személynevek szerint is szelektálhatók, illetve olyan textológiai jelenségekre is szűrhetők, mint az idegenkezűség vagy írószöveg-váltás.

4. A szöveg mint hálózat

A keresések, lekérdezések által számszerűsített adatokat kapunk a platformon, amelyek lehetővé teszik a szöveg adatok felőli, statisztikai alapú megközelítését. Debreczeni Attila szerint az a tény, hogy a kiadás „adatbázisra épül, amelyből a felhasználói kéréseknek megfelelően algoritmusok alapján időről időre generálódik, azt is jelenti, hogy nincs állandó lineáris szerkezete, mint a nyomtatott médiumban létező kiadásoknak. Az adatbázisban tárolt szövegek mint információs

"Minden kiadáshoz különböző szűrési szempontok állíthatók be a dHUpla felületén, amelyek az adott szövegtörzs TEI XML-alapú, metanyelvi kódolásán alapulnak."

alapegységek egymással hálózatos viszonyban vannak [...]. E viszonyokban a szövegek filológiai jellemzői testesülnek meg, s válnak megjeleníthetővé a kiadási szerkezetben, funkciókban kódolt lekérdezési lehetőségek szerint. [...] Az így felfogott elektronikus kiadás nem könnyebben használható könyv immár, hanem olyan új kutatási lehetőségek tárháza, amelyek a lineárisan rögzített nyomtatott médiumban elképzelhetetlenek" (Debreczeni 2020, p. 406–407). Ez a megfogalmazás ugyanígy jellemző lehet a forráskiadásokra is, attól függően, hogy milyen adatokat azonosítunk be szövegükben. Így a digitális kiadás egyfajta adatbázisként is működik, amely hozzájárul az összefüggések, mintázatok feltáráshoz, kontextusba helyezéséhez (Schöch 2013), ami nagy méretű szövegtörzsek értelmezhetősége, befogadhatóvá tétele szempontjából különösen hasznos.

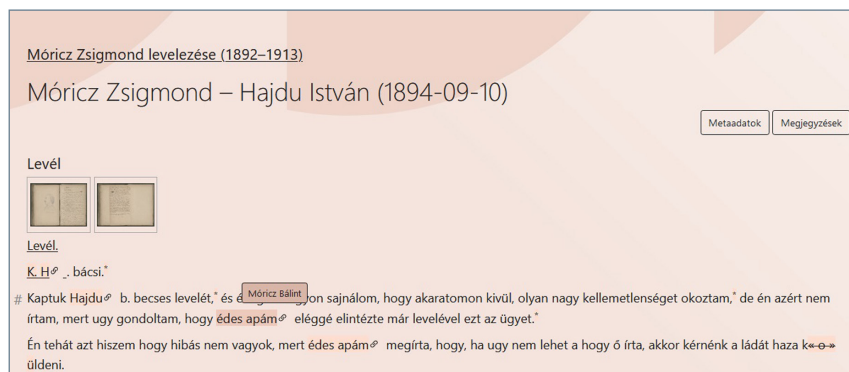
A lineáris szerkezet felbomlásával megváltozik a szöveghez való viszonyulásunk. Elektronikus kiadásnál „lehetőség nyílik arra, hogy egy-egy levélnek a lineáris egységen belül létező elemeit elkülönítsük, ezeket összekapcsoljuk más levelek tartalmilag érintkező szövegelemeivel, s egységes magyarázattal ellátva együtt olvashatóvá tegyük azokat" (Debreczeni 2020, p. 399). Az adatok összefüggésének, hálózatoságának megmutatása ezzel szemben a papíralapú kiadásnál „elsősorban jegyzetekben valósítható meg, de csak bonyolult ide-oda utalásokkal, rendkívül nehézkesen és kevésbé átláthatóan" (Debreczeni 2020, p. 399). Egy olyan kiterjedésű szövegtörzs esetén, mint a Móricz-levelezés, egy papíralapú kiadásnál az összefüggések közötti eligazodásban hagyományosan a jegyzetek, a névmutatók és a tartalomjegyzék tudnák navigálni az olvasót. A digitális szövegkiadásban ezzel szemben a feltárt tartalmi összefüggések megjelenítése jegyzetekkel összekötött levélutalásokkal történik, ami a levelek közötti egyszerű átjárást teszi lehetővé, így a tematikus összefüggések mentén haladva is olvashatjuk a szöveget. A levélszövegek külső hivatkozásokkal történő összekapcsolása egy másik szinten is megteremt a hálózatoságot, így integrálva a szövegkiadást egy kiterjedtebb, tematikus közegben. A digitális szöveg egyrészt lineárisan is olvasható, másrészt a kéziratkép összekapcsolása az átírással további lehetőséget ad a szöveg mélyebb vizsgálatára. Mindez nem jelenti azt, hogy az elektronikus kiadások kizárják a papíralapú kiadások létjogosultságát. A tendenciák inkább azt mutatják, hogy a kétféle médium nagyon is tud egymás mellett létezni: „épp különeműségük miatt másra jó, másként használható egy kritikai kiadás papíron és másra jó, másként használható képernyőn: mindkettőre szükség van" (Debreczeni 2020, p. 401).

5. A változandóság előnyei és hátrányai

A digitális kiadások sokat emlegetett jellemzője a rugalmasság, ami egyben jelent előnyt és hátrányt is. Egy digitális kiadás bármikor javítható, adatokkal gazdagítható, összeköthető külső hivatkozásokkal, új funkciókkal továbbfejleszhető, így mindig naprakész és innovatív verzió állhat a rendelkezésünkre, terjedelemtáron kívül. A dHUplán megjelent eddigi kiadásokra is jórészt a folyamatosság állapota jellemző: a Móricz-kiadásban a kritikai jegyzetek és a külső névterekkel való összekötés javításán dolgozunk, Kiss József levelezését a PIM-es és az OSZK-ban fellelhető levelek kiadásával bővítjük, és a TEI XML-annotációk, -jelölések dHUplán való megjelenítésének fejlesztése is folyamatban van. Mindez hozzájárul ahhoz, hogy a felhasználók és a kutatók igényeit kiszolgáló, egyre gazdagabb, naprakész és innovatív publikációk álljanak rendelkezésre. Ahhoz viszont, hogy a folyamatos bővülés és fejlesztés megvalósuljon, rugalmas infrastruktúrára és verziókövető rendszerekre van szükség, valamint egy digitális infrastruktúra fenntartása állandó informatikai támogatást is igényel, ami sokszor kihívást jelentő feladat a köztisztviselők számára.

"A dHUplá-
infrastruktúra
közeppontjában
álló »git« alapú
rendszer előnye
a visszakövet-
hetőségen túl,
hogy teljesen
szoftver-
független..."

Maróthy Szilvia kiemeli: „A hálózati kiadás tartalma akár pillanatról pillanatra változhat, de nem állandók a létrehozók és üzemeltetők sem, s az elérés is gyakran módosul. A tudományosság egyik alapfeltétele, a visszakereshetőség látszik sérülni” (Maróthy 2017, p. 299). Hazai és nemzetközi kiadásoknál egyaránt belebotlunk abba a jelenségbe, hogy néhány évvel korábbi publikációk már nem érhetőek el, vagy változott az elérésük, esetleg félbemaradt a fejlesztésük.¹⁵ Így a változékonyság – pozitív aspektusai mellett – az instabilitás, a konstans befejezetlenség benyomását is keltheti a digitális kiadások esetében, ami szemben áll a hagyományos, papíralapú publikálás megbízhatóságával, statikusságával.



5. ábra: minta a beépített névmutató működésére (Forrás: Mórícz Zsigmond levele Hajdu Istvánnak)¹¹

A dHUpa-infrastruktúra középpontjában álló „git” alapú rendszer előnye a visszakovethetőségen túl, hogy teljesen szoftverfüggetlen, így a közgyűteményi szféra különböző infrastruktúráival könnyen kompatibilissé tehető (Mihály 2022, p. 347–351). Bár a visszakovethetőség problémája ebben az esetben sem védhető ki teljesen egy szolgáltatás megszűnése esetén, azonban nagy rugalmasságot jelent, hogy a rendszer adatvesztés nélkül újra felépíthető másik szoftverkörnyezetben.

A digitális kiadásoknál fennáll a tartalmi elévülés veszélye a „halott linkek” kezelésében, amiben egyik fontos jellemzőjük, a hálózatoság sérül. A hivatkozhatóság azonban állandó azonosítók használatával is lehetséges. A dHUpán levő kiadásoknál például külső névterekre állandó azonosítókkal és nem magára a névtéroldalra hivatkozunk. Így, ha megváltozik egy weboldal vagy adatbázis elérési útja, az állandó azonosítók továbbra is elvezetnek a célhelyre.

Ha egy kiadást sosem tekintünk lezárt egészként, akkor olyan releváns kérdések vetődnek fel, mint hogy: „miben is áll egy ilyen kiadás önazonos volta, mikortól kiadás egy kiadás, illetve a változások által nem lesz-e teljesen más, mint addig? Mire is hivatkozunk egyáltalán ebben az esetben?” (Debreczeni 2020, p. 410) Nem csak a hivatkozhatóság jelent azonban problémát. Állandóan felmerülnek arra vonatkozó kérdések, hogy egy feldolgozott digitális szövegtörzs mikortól vehető publikációnak. „Azt tekintjük-e publikációnak, melyet a létrehozó (s egyben kiadó) publikációként határoz meg? De mit jelent a publikáció, a nyilvánosság tétele fogalma a hálózaton, mikor a lehetősége elvben mindenki számára adott, s élünk is vele nap mint nap?” (Maróthy 2017, p. 306) Ezek a kérdések arra is rávilágítanak, hogy a digitális kiadások népszerűsítésének nincs olyan kialakult kulturális gyakorlata, mint a papíralapú publikációknak (pl. bemutatók, felolvasások). A tapasztalatok azt mutatják, hogy a láthatóság és befogadhatóvá tétel érdekében a digitális kiadásoknál is szükséges olyan közeg kialakítása, ahol ez a gyakorlat megvalósul, még akkor is, ha a szabad elérhetőség eredendően mindenki számára adott.

"A kihívások ellenére a digitális szövegkiadások rengeteg új lehetőséget nyújtanak a tartalom-szolgáltatásban és a kutatás-támogatásban."

A kihívások ellenére a digitális szövegkiadások rengeteg új lehetőséget nyújtanak a tartalomszolgáltatásban és a kutatástámogatásban. Adatbázisjellegükből adódóan hozzájárulnak olyan összefüggések, mintázatok feltárásához, amelyek a papíralapú kiadásoknál kevésbé láthatóak, és amelyek új szempontokkal egészíthetik ki a hagyományos kutatásokat. A dHUpla keretrendszere egy működő infrastruktúrát biztosít kéziratok közreadására, az intézményi kereteken túl külső kutatók és intézmények számára.

Lábjegyzet

- ¹ A digitális szövegkiadások projektvezetője Mihály Eszter, a DBK igazgatóhelyettese és a Digitális Filológiai és Webarchiválási Osztály vezetője.
- ² Transkribus szoftver, Elérhető: <https://readcoop.org/> (Utolsó elérés: 2024.03.15.)
- ³ TEI – Text Encoding Initiative nemzetközi szabvány, Elérhető: <https://tei-c.org/> (Utolsó elérés: 2024.03.15.)
- ⁴ Oxygen XML Editor program, Elérhető: <https://www.oxygenxml.com/> (Utolsó elérés: 2024.03.15.)
- ⁵ Petőfi Sándor levelezése, Elérhető: https://dhupla.hu/text/PL-fond-VII-0083-0002_tei (Utolsó elérés: 2024.03.15.)
- ⁶ Külön specifikáció készül a levél, a monografikus (pl. napló) és az analitikus (pl. folyóirat) -típusú szövegtörzsek számára. A specifikációt hamarosan publikussá tesszük a dHUpla felületén.
- ⁷ GitLab rendszer, Elérhető: <https://about.gitlab.com/> (Utolsó elérés: 2024.03.15.)
- ⁸ Kiss József levelezése, Elérhető: <https://dhupla.hu/collection/kiss-jozsef-levelezes> (Utolsó elérés: 2024.04.15.)
- ⁹ Vershaza, Elérhető: <https://dhupla.hu/collection/vershaza> (Utolsó elérés: 2024.04.15.)
- ¹⁰ Vörösmarty Mihály: Szózat fogalmazvány, Elérhető: https://dhupla.hu/text/Szozat_fogalmazvany_kezirat_mta1/text-image#facs_1 (Utolsó elérés: 2024.04.15.)
- ¹¹ Móricz Zsigmond levelezése, Elérhető: <https://dhupla.hu/collection/moricz-zsigmond-levelezes> (Utolsó elérés: 2024.04.15.)
- ¹² A Móricz-kiadásról bővebben lásd: Cséve A. (2022) Előszó, In: Móricz Zsigmond levelezése (1892–1913) Elérhető: https://dhupla.hu/text/mzs-lev_eloszo (Utolsó elérés: 2024.05.12.)
- ¹³ Teleki László tanácsadása a nevelésről, Elérhető: <https://dhupla.hu/collection/teleki-laszlo-tanacsadasa> (Utolsó elérés: 2024.03.15.)
- ¹⁴ Fábri A. (2022) Arany idők – Esmék és emberek gróf Teleki László neveltetésében, Kortárs Könyvkiadó, Budapest, ISBN: 9789634351085
- ¹⁵ Saját példaként szolgál a Móricz-levelezéskiadás, amelynek első verziója a PIM és a DBK ITI együttműködése alapján 2019-ben a DigiPhil keretrendszerében készült el. A 2.0 verzió a PIM Digitális Bölcsészeti Központ platformján, a dHUplán vált publikussá, ami 2021 decemberétől az OSZK DBK saját platformjaként szolgáltatja tovább a levelezéskiadást. Bővebben lásd: Cséve, A. (2022) Előszó. In: Móricz Zsigmond levelezése (1892–1913). Elérhető: https://dhupla.hu/text/mzs-lev_eloszo (Utolsó elérés: 2024.03.13.)

Felhasznált és ajánlott irodalom

Bia, A. (2023) *On Long Term Digital Preservation and Document Maintenance, and the Role of XML and TEI*, In: Annisus, M., Arestau, E., Burkhardt, J., Herold, N., Sierig, R. (szerk.) *Vielfalt und Integration - diversità ed integrazione - diversité et intégration: Sprache(n) in sozialen und digitalen Räumen, Eine Festschrift für Elisabeth Burr*, [online] p. 283–295, Universität Leipzig, Lipcse, Elérhető: <https://ul.qucosa.de/api/qucosa%3A84926/attachment/ATT-0/> (Utolsó elérés: 2024. 06. 13.)

- Cséve, A. (2022) *Előszó*, Elérhető: https://dhupla.hu/text/mzs-lev_eloszo (Utolsó elérés: 2024.03.13.)
- Debreczeni, A. (2014) *Kritikai kiadás papíron és képernyőn*, In: Czifra, M., Szilágyi M., (szerk.) *Textológia, filológia, értelmezés: Klasszikus magyar irodalom*, p. 26–42, Debreceni Egyetemi Kiadó, Debrecen, ISBN 978-963-318-468-4
- Debreczeni, A. (2020) *Az elektronikus kiadás színe és visszája*, [online] In: *A humán tudományok és a gépi intelligencia*, p. 48–61, Gondolat Kiadó, Budapest, ISBN 978-963-6938-70-3 Elérhető: <https://real.mtak.hu/88668/> (Utolsó elérés: 2024. 05. 17)
- Maróthy, Sz. (2017) *Elektronikus szövegkiadások a könyvtárban*, *Tudományos és Műszaki Tájékoztatás*, 64(6), p. 298–309. Elérhető: <https://journals.bme.hu/tmt/article/view/34585/20298> (Utolsó elérés: 2024. 05. 17)
- Maróthy, Sz. (2020) *Digitális bölcsészet*, Budapest, Akadémiai Kiadó, Elérhető: https://mersz.hu/dokumentum/m794aszad__1/ (Utolsó elérés: 2024. 06. 13.)
- Mihály, E. (2022) *Mi az a dHUpLa? A Digitális Bölcsészeti Platform bemutatása, Valós térben – Az online térért*, In: Hungarnet Egyesület, (2022) *Valós térben - Az online térért: Networkshop 31: országos konferencia*, Debreceni Egyetem, Debrecen 2022. április 20–22, Budapest, p. 345–358, ISBN 978-615-82243-0-7 <https://doi.org/10.31915/NWS.2022.44>
- Palkó, G. (2016) *Mit jelent a digitális filológia a szemantikus web korában?*, *Magyar Tudomány* 177(11), p. 1316–1322, Elérhető: https://epa.oszk.hu/00600/00691/00158/pdf/EPA00691_mtud_2016_11_1316-1322.pdf (Utolsó elérés: 2024. 06. 13.)
- Schöch, C. (2013) *Big? Smart? Clean? Messy? Data in the Humanities, The Dragonfly's Gaze*, Elérhető: <https://dragonfly.hypotheses.org/443> (Utolsó elérés: 2024. 06. 13.)
- Szűcs, K. Á. (2021) *Automatikus kézírás-felismeretetés Kiss József levelezésén*, In: Networkshop 30: országos online konferencia, Eötvös Loránd Tudományegyetem, Budapest, 2021. április 6–9, p. 330–337, <https://doi.org/10.31915/NWS.2021.8>
- Thomas, S. (2020) *«A digital edition is not visible» – some thoughts on the nature and persistence of digital editions*, In: *Zeitschrift für digitale Geisteswissenschaften*, Wolfenbüttel, https://doi.org/10.17175/2020_005, Elérhető: https://zfdg.de/sites/default/files/pdf_uploads/digital_editions_2020_0.pdf (Utolsó elérés: 2024. 06. 13.)

