

Kötetünk olvasói körében a könyvtár: valóság. Fontos helyszíne életünknek hallgatóként, egyetemi oktatóként, a társadalom bármely korosztályának tagjaként vagy a könyvtárak munkatársaiként. Annak a valóságnak szerves része, amely a tanulás, a kutatás háttereként, hivatásunk céljaként, de akár közösségi térként, a kulturális örökség őrzőjeként vagy a szabadidő hasznos eltöltésének segítőjeként értelmet ad mindennapjainknak.

Reményeink szerint a kötetben szereplő tudományos közlemények elegendő alapot adnak arra, hogy a könyvtári valóság újdonságait, lehetőségeit és kérdéseit szakmai vagy laikus olvasóként egyaránt megismerjük. Ezáltal mindannyian pontosabban láthatjuk majd, mire képesek a könyvtárak a XXI. században és mivel szolgálhatják a könyvtárosok és könyvtártudósok az emberiséget.

ISBN 978-963-489-017-1



9 789634 890171



VKIV
VALÓSÁGOS KÖNYVTÁR – KÖNYVTÁRI VALÓSÁG
KÖNYVTÁR- ÉS INFORMÁCIÓTUDOMÁNYI TANULMÁNYOK 2018



VALÓSÁGOS KÖNYVTÁR – KÖNYVTÁRI VALÓSÁG

KÖNYVTÁR- ÉS INFORMÁCIÓTUDOMÁNYI
TANULMÁNYOK 2018



Szerkesztette
KISZL PÉTER
és CSÍK TIBOR

A TUDOMÁNYOS ELEKTRONIKUS SZÖVEGKIADÁSOK FELDOLGOZÁSA*

MARÓTHY SZILVIA

ELTE BTK ITDI Magyar Barokk Doktori Program, doktorjelölt

TARTALMI ÖSSZEFOGLALÓ

A digitális dokumentumok speciális típusa a tudományos elektronikus szövegkiadás. A számítógépes filológia, számítógépes szövegfeldolgozás ugyan a XX. század közepéig nyúlik vissza, de a weben elérhető szöveges tartalmaknak csak töredékét képezik a kritikai igényrel készített és közreadott szövegkiadások. Hazánkban a kilencvenes évek elejétől fogva készülnek ilyen összeállítások, az elmúlt huszonöt év termése azonban szinte láthatatlan: a könyvtári nyilvántartások alig tájékoztatnak róluk, gondozásukat a pályázati finanszírozás hullámvölgyei teszik lehetetlenné. Kutatási projektemnek az a célja, hogy e kiadásokat a könyvtári szabványnak megfelelően feldolgozzam, s ezáltal láthatóbbá tegyem, valamint, hogy kaput nyissak egy esetleges archiválási projekt számára.

Bevezetés

A tartalmilag és technikailag összetett tudományos elektronikus szövegkiadások (pl. kritikai vagy forráskiadások) csak ritka esetben állnak kapcsolatban a könyvtári rendszerekkel. Gyűjtésükkel, feldolgozásukkal a könyvtárak kevésbé foglalkoznak, így a katalógusokban is alig akadhatunk a nyomukra. A kiadások szolgáltatását a készítői saját feladatuknak tekintik, a könyvtár pedig távoli hozzáférésüként regisztrálja azokat.

A digitális filológiai közösség két nemzetközi katalógust tart számon, ám e két katalógus csupán egyetlen magyar tételt tartalmaz, ezért jelenleg a magyar nyelvű, illetve magyar vonatkozású kiadások katalógusán dolgozom az ELTE BTK Digitális Bölcsészeti Központtal együttműködésben, az ÚNKP-17-3 jelű ösztöndíj keretében. Az alábbiakban a tudományos elektronikus kiadás definícióival, egy kiadás verzióinak ismertetésével, valamint a kiadástípus könyvtári szabványoknak megfelelő feldolgozásával foglalkozom.

A tudományos elektronikus szövegkiadás fogalma

Az általam használt *tudományos elektronikus szövegkiadás* részint az angolszász szakirodalomban állandósult *digital scholarly edition* fogalomra, részint a hazai szakirodalomban használt kifejezésekre alapoz.¹ Előbbinek számos definícióját adja *Patrick Sable*. Röviden összefog-

* A tanulmány az Emberi Erőforrások Minisztériuma ÚNKP-17-3 kódszámú Új Nemzeti Kiválóság Programjának támogatásával készült. Köszönöm Golden Dánielnek, Parádi Andreának és Vadai Istvánnak, hogy a hazai számítógépes filológia hőskorszakáról, annak terméséről értékes adalékokkal szolgáltak.

lalva: a tudományos kiadás a történeti dokumentum kritikai reprezentációja. Három kritériuma van: 1) létező forrásokat, történeti dokumentumokat dolgoz fel; 2) forrást reprezentál és azt metaadatokkal látja el; 3) szövegkritikai munkát végez a forrás reprezentálásakor.²

Nem tartoznak tehát a kiadástípus alá azon gépi vagy kézi eszközökkel digitalizált szövegek, melyek nem reflektálnak a szöveg és hordozója sajátosságaira. Sahle további megszorításokkal él a *digital* jelző kiérdeklése kapcsán: a digitalizált papírkiadás nem, csak az eleve digitális (*born digital*) kiadások tartoznak ide. Ez a határ azonban az egyre kifinomultabb digitalizálási projekteket tekintve elmosódni látszik: például a DigiPhil projekt által elindított Mikszáth kritikai kiadássorozat,³ mely a papíralapú kiadást szemantikus gazdagított TEI XML formátumban publikálja a weben, ne lenne elektronikus szövegkiadás? Véleményem szerint az. Ezért saját gyűjtésem során nem alkalmazom Sahle azon megkötését, mely szerint a nem digitálisan született kritikai kiadások az elektronikus szövegkiadás típusán kívül esnek.

A hazai diskurzusban jelenleg az MTA Textológiai Munkabizottságának állásfoglalása lehetne mérvadó, ám sajnálatos módon – bár maguk a tagok is sürgetik – ez 2004 óta nem frissült.⁴ A dokumentum szerint az elektronikus kiadás a szövegkiadások egy típusa a kritikai, genetikus stb. kiadások mellett. Alapvető jellemzőjeként a rendelkezésre álló szövegforrások teljesség igényével történő feldolgozását nevezi meg, ám ez inkább a klasszikus értelemben vett kritikai kiadás, vagy más esetben a forráskiadás jellemzője. Hiszen az elektronikus kiadás szabadsága éppen azt jelenti, hogy ezen típusok bármelyike, vagy mindegyike lehet.⁵

A tudományos és az elektronikus kifejezéseken túl a *kiadás* is magyarázatra szorul. Mint az alábbi példa mutatja, a megjelenési közeg sajátossága, hogy nem feltétlenül tudunk kézzelfogható, megtapasztalható példányokra rámutatni. Az SGML, majd az XML használatával pedig a rögzített tartalom és annak megjelenítése is markánsan szétvált, így a klasszikus *kiadás* fogalom már nem alkalmazható.

A Balassi-életmű hálózati kiadása

A szövegkiadások feldolgozása során felmerülő kérdéseket a legkorábbi hálózati kritikai kiadáson szeretném bemutatni. Az egyik változatában a weben jelenleg is elérhető *Balassi Bálint összes verse, hálózati kritikai kiadás* kezdetei 1993-ig nyúlnak a visszaemlékezések szerint. Abban az évben készült el a Balassa-kódex (OSZK Quart. Hung. 3247) betűhű forráskiadása Gopher⁶ alatt, melyet akkor még maguk a szöveg gondozói sem láthattak működésben.⁷

Ezt követően két CD-ROM-os kiadvány készült a Balassa-kódexből, egy 1994–95-ben, mely a fentebbi grafikus változata, majd még egy 1996-ban, mely a grafikus böngészőben elérhető változatot őrizte meg.⁸ Ezeket Tóth Tünde 2004-ben két tanulmányában ismerteti, így most nem térek ki rá.⁹ A szerző megjegyzése szerint az 1996-os online kiadvány 1998-ban megsemmisült. Az Internet Archive azonban több mentést

is készített róla, melyek a Wayback Machine szolgáltatásnak köszönhetően elérhetők. Érdemes lenne lépéseket tenni ennek, és az ehhez hasonlóknak könyvtári megőrzésére is a webarchiválási projekt keretében.

1998-ban készült el a Balassi-életmű teljességre törekvő kritikai kiadása, mely a Balassa-kódexen felül számos kéziratos és nyomtatott forrást dolgozott fel, és kritikai szöveget és forrásközlést is adott. Bár a készítőik figyelemmel követték a nemzetközi eseményeket: a TEI szabvány alakulását, az SGML jelölőnyelv előnyeit, melyben a TEI is kifejezte ajánlását, illetve az akkor bevezetésre kerülő XML-t, a kiadás mégis HTML-ben készült el. A struktúra és a megjelenítés szétválasztása, mely az SGML és az XML nagy előnye, a kritikai kiadások készítésében is alapvető igény. A készítőik fő célja azonban a webes közzététel, a széles körű hozzáférés biztosítása volt, s ez az akkori lehetőség mellett egyértelműen a HTML választását vonta maga után.¹⁰

2002-ben tovább fejlesztik, bővítik a kiadást. Az előszóban a korábbi verziókról röviden megemlékeznek a készítőik. Ebből tudhatjuk meg, hogy volt egy köztes, 1999-es változat is. Sajnos azt, mi a különbség a három verzió között, az előszóból nem tudjuk meg.

„A jelen kritikai kiadás 0. változata (1993) csak a Balassa-kódex átírását, fényképhasonmását és a jegyzetek egy részét tartalmazta. Az 1.0 változat (1998), az 1.1. változat (1999), valamint a 2.0 változat (2002) tartalmazza Balassi Bálint teljes költői életművét, a megállapított kritikai szöveget, annak átírását, a forrásszövegek betűhű átírásait és fényképhasonmásait, a költemények részben újonnan megállapított sorrendjét, a Répertoire de la poésie hongroise ancienne (RPHA, ©1979-2001) kivonatait, a szövegkritikai jegyzeteket és a szövegmagyarázatokat.”¹¹

A verziók közötti különbség valószínűleg főként a digitális hasonmások számában keresendő, erre utal Tóth Tünde idézett tanulmányának egyik lábjegyzete: „Technikai okokból nem tudtuk az összes forrás fakszimiléjét egyszerre beszerezni. Így a kiadás folyamatosan bővülő verziói egyre gazdagabbak.”¹² Ha azonban komolyan vesszük a verziók szintjén történő feldolgozást, itt azt kell mondanunk, hogy az 1998-as és 1999-es verziók feltehetőleg elvesztek (vagy felülírták a fájlokat), vagy a készítőik saját adathordozóin hevernek. Változtatásokat könyvelő dokumentációt nem találtam. Jelenleg az ELTE szerveréről a 2002-es verzió érhető el.¹³

2004-ben egy jelentősebb tartalmi és technikai frissítésen esett át a kiadás, amely az újabb nyomtatott kiadások digitális fakszimiléivel bővült, Unicode karakterkódolásra tértek át, a felmerült tartalmi hibákat javították. Technikai szempontból a legfontosabb újítás az akkor nemrégiben kiadott, már XML alapú TEI P4 ajánlásra való áttérés volt. Ezt részint automatikus, részint kézi konverziókkal érték el. Az így elkészült XML fájlokból generáltak HTML megjelenítéseket, s ezen felül új képezelőt terveztek a kiadáshoz. Sajnos ez a 3.0 verziószámú kiadás weben jelenleg nem elérhető, bár a forrásfájlok javarészt megvannak a készítőik adathordozóin több példányban – érdemes lenne ezeket is archiválni.

2016-ban az Országos Széchényi Könyvtár (OSZK) rövid időre felkarolt néhány hálózati szövegkiadást. Ennek keretében *Parádi Andrea* a 2004-es TEI P4 szabványú

kódolást P5-re frissítette. Ez szintén csak részben automatikus konverziót jelentett, tehát az áttérés szakértői munkát is igénybe vett. Fontos strukturális változás továbbá, hogy míg a korábbi kiadásokban a versek feldolgozása külön fájlba került, itt egy fájl tartalmazza a teljes kiadást, s ennek megfelelően lettek a belső hivatkozások is kialakítva. Megjelenítés ehhez még nem készült, kérdés tehát, tekinthetjük-e ezt egyáltalán kiadásnak? A fájl elhelyezése sem megoldott: bár az OSZK megrendelésére készült, tudomásom szerint annak semmilyen archívuma nem őrzi, így szintén csak magántulajdonban lévő adathordozókon létezik.

Kutatási terv

A fenti példából látszik, hogy már a dokumentumok leírása több problémát vet fel. Nem pusztán a klasszikus dokumentum fogalom okozza a nehézségeket, mely a kiadást térben és időben lezárt, végleges egységként képzei el, hanem az is, hogy a különféle verziók dokumentálására és megőrzésére a létrehozók, megrendelők nem fordítottak különösebb figyelmet. Így mind a leírni kívánt példány(ok) előkerítése, mind pontos leírása kutatómunkát igényel.

Bár nem jelentős mennyiségű rekordról van szó, ezért a választásnak kevésbé nagy a tétje, mégis megfontolandó, a kiadványok leírásához a bevett MARC leírást vagy valamely entitáskapcsolatokon alapuló korszerűbb könyvtári szabványt alkalmazzam-e. Az elektronikus szövegkiadásokra – mint digitális objektumokra általában – az utóbbit tartom alkalmasabbnak. A Balassi-versek hálózati kiadásának kilenc verziója ismeretes: ezek mind tartalmukban, mind technikai megvalósításukban különböznek. Azonban mégis egy kiadványnak újabb és újabb bővített kiadásai, mely tényre a verziószámokkal és az előszókban egyaránt felhívják a szerzők a figyelmet. Ezt a viszonyt jól kezeli mind az FRBR, mind a BIBFRAME modell. Jelentős különbség a kettő között abban látszik, hogy míg előbbi a dokumentumok és jellemzőik kapcsolatát hierarchikusan, utóbbi hálózatosan képzei el.

A könyvtári szabványnak megfelelő leírás mellett azonban egy egyedi bibliográfia struktúrájához alkalmazkodó leírás elkészítésére is szükség van. A tudományos elektronikus szövegkiadások kutatásával foglalkozó *Greta Franzini* kollégáival nemzetközi nyilvántartás összeállítását kezdeményezte. Ez a katalógus a szigorúbb értelemben vett tudományos elektronikus szövegkiadásokat gyűjti egybe. Ennek meghatározása a fent idézett Patrick Sahle definícióit veszi alapul: filológiai munkával előállított, eleve digitálisan készült szövegkiadásokat jelent. A projekt folyamatban van, az összegyűjtött tételeket még nem mind dolgozták fel. Emellett nyitott, ahhoz bárki csatlakozhat, a GitHub felületén keresztül, ahol a katalógus elérhető és letölthető.¹⁴ Nyilvántartásuk elsődleges célja, hogy összegyűjtsék és elemezzék, milyen elektronikus szövegkiadások készültek eddig, mik a jellemző gyakorlatok, típusok. Szempontrendszerük igen sokrétű, az alapadatokon felül például a technikai részletek, az intézményes és pénzügyi háttér is helyet

kap. E nemzetközi gyűjtést mindenképpen szeretném bővíteni az általam feldolgozott tételekkel, hiszen a készítőknél kevés rálátásuk van a magyarországi anyagra.

Összefoglalás

A tudományos elektronikus szövegkiadások feldolgozásának célja, hogy azok megtalálhatóvá, felhasználhatóvá és hivatkozhatóvá váljanak. A hálózaton szolgáltatott dokumentumokkal a klasszikus értelemben vett katalóguson rég túl kellett lépjen a könyvtár, így nem kérdés, hogy e néhány tétellel is érdemes bővíteni azokat. A kiadások gyűjtésével és leírásával nem titkolt célom továbbá, hogy azok megőrzését is elősegítsem, szorgalmazzam. Ez azonban csak intézményi infrastruktúra keretein belül működhet: láthattuk, a szöveggondozók efemer adathordozóin tárolt példányokhoz már alig férhetünk hozzá.

Jegyzetek

1. A kiadástípussal kapcsolatos terminológiai álláspontokat, vitákat egy korábbi tanulmányomban részletesen ismertettem, így e helyet nem térek ki rá. MARÓTHY Szilvia: Elektronikus szövegkiadások a könyvtárban. = Tudományos és Műszaki Tájékoztatás, 64. évf. 2017. 6. sz. 298–309. p. Forrás: <http://tmt.omikk.bme.hu/tmt/article/view/1000> [2018. január 22.]
2. SAHLE, Patrick: About a catalog of: Digital Scholarly Editions, v 3.0, snapshot 2008ff. Utolsó módosítás: 2017. március 22. Forrás: <http://www.digitale-edition.de/vlet-about.html> [2018. január 22.] Lásd még SAHLE, Patrick: What is a Scholarly Digital Edition? In: Digital Scholarly Editing: Theories and Practices. (Eds.) Driscoll Matthew James, Pierazzo, Elena. Digital Humanities Series, 4. Open Book Publishers, 2016. 19–40. p. Forrás: <http://www.openbookpublishers.com/product/483/digital-scholarly-editing-theories-and-practices> [2018. január 22.]
3. DigiPhil: Mikszáth Kálmán művei. Forrás: <http://www.digiphil.hu> [2018. január 22.]
4. MTA Textológiai Munkabizottság: Alapelvek az irodalmi szövegek tudományos kiadásához. 2004. <http://textologia.iti.mta.hu/alapelvek.pdf> [2018. január 22.] Az alapelvek megújítását legutóbb Debreczeni Attila sürgette *A humán tudományok és a gépi intelligencia* c. konferencián 2017. november 20-án elhangzott Az elektronikus kiadás színe és a visszaja c. előadásában.
5. Lásd erről részletesen: MARÓTHY i. m. 298–299. p.
6. A World Wide Webet megelőző internetes protokoll a kilencvenes évek elején.
7. TÓTH Tünde: Online kritikai szövegkiadás Magyarországon az ezredfordulón. = Helikon, 3. sz. 2004. 417–440. p., 426. p.; és HORVÁTH Iván: Számítógép és irodalomtudomány. = Alföld, 3. sz. 2010. 9–28. p. 14. p.
8. Horváth Iván (szerk.), Vadai István (átírás), Szedmák Sándor (képek), Kovács István (program): Balassa-kódex. Budapest, ELTE BTK CHER–ELTE BTK Informatikai Osztály–Hypermedia Systems, 1994–95.; Horváth Iván (szerk.), Vadai István (átírás), Szedmák Sándor (képek), Kovács István (program): Balassa-kódex (tanulmányokkal). Budapest, ELTE

MARÓTHY SZILVIA

- BTK CHER–Hypermedia Systems–IDG, 1996. <http://www.idg.hu/expo/oktigon/balassa> [2018. január 22.]
9. TÓTH 2004a i. m.; TÓTH Tünde: Balassi Bálint verseinek új, hálózati kritikai kiadása. = *Hungarologische Beiträge*, 15. 2004. 79–84. p.
 10. TURI László – TÓTH Tünde – GOLDEN Dániel: Virtuális örökkévalóság: objektumok a digitális könyvtárban. *Palimpszeszt*, 10. 1998. Forrás: <http://magyar-irodalom.elte.hu/vita> [2018. január 22.]
 11. HORVÁTH Iván – TÓTH Tünde: Névjegy. In: H. I., T. T.: Balassi Bálint összes verse, hálózati kritikai kiadás. Forrás: <http://magyar-irodalom.elte.hu/gepesk/bbom/inevjegy.htm> [2018. január 22.]
 12. TÓTH 2004b i. m., 82. és 84. p.
 13. Horváth Iván – Tóth Tünde – Vadai István (szerk.): Balassi Bálint összes verse, hálózati kritikai kiadás, v. 2.0. Gépeskönyv, 2002. <http://magyar-irodalom.elte.hu/gepesk/bbom> [2018. január 22.]
 14. FRANZINI, Greta: Catalogue of Digital Editions. GitHub. https://github.com/gfranzini/digEds_cat [2018. január 22.] Lásd még: FRANZINI, Greta, TERRAS, Melissa, MAHONY, Simon: 9. A Catalogue of Digital Editions. In: DRISCOLL, PIERAZZO i. m. 161–182. p. <http://www.openbookpublishers.com/product/483> [2018. január 22.]

Maróthy Szilvia az ELTE BTK Irodalomtudományi Doktori Iskolájának doktorjelöltje. Az ELTE ÚNKP ösztöndíjasaként a hazai tudományos elektronikus kiadásokat dolgozza fel. A Digitális Bölcsészet folyóirat szerkesztőjeként a hazai kutatások népszerűsítésével, a műhelyek közötti párbeszéd ösztönzésével foglalkozik.