

Maróthy Szilvia: A nyílt és a zárt tudományról

A tudományos tevékenység feltételrendszere és etikai normái szerint a tudományos eredményt akkor tudjuk elismerni, ha 1) hozzáférünk, 2) azon forrásokhoz is hozzáférünk, melyeken az eredmény alapszik, 3) az előbbi kettő birtokában ellenőrizni, illetve reprodukálni tudjuk azt. Tanulmányomban azt vizsgálom, ezek a feltételek mennyiben teljesülnek a webes közegben, mely az utóbbi évtizedben a tudományos publikálásnak is a legjelentősebb felülete lett. A webes technológiák elméletileg sokkal több lehetőséget kínálnak az említett feltételek teljesítésére, valamint a tudásmegosztásra, a kollaborációra. Mást mutat-e a gyakorlat? Ezt vizsgálom alapvetően bölcsészettudományi példákon keresztül.

A nyílt tudomány (*open science*) megnevezés azon elvek összességét takarja, melyek a tudás szabad hozzáférését és újrafelhasználását támogatják. Ilyen a tudományos publikálásban is sokat emlegetett nyílt hozzáférés (*open access*), mely az elmúlt évtizedben jelentős intézményesülésen ment keresztül. Inkább mozgalom jellegű, másfelől piaci modellt képvisel a nyílt forráskód (*open source*), mely elsősorban a szoftverek, másodsorban a digitális tartalom számítógépes kódjainak, fájljainak megosztását, közösségi fejlesztését és felhasználhatóvá tételét takarja (lásd még a nyílt adat, vagy *open data* fogalmát). A magyar nyelvű információs portál megfogalmazása:

Az Open Science olyan tudományos gyakorlat, ahol lehetőség nyílik az együttműködésre, a kutatási adatok, laboratóriumi jegyzőkönyvek és egyéb, a kutatás folyamatában keletkezett információk szabad felhasználására, engedélyezve az újrahasznosítást, a reprodukálást, a terjesztést. Egy olyan mozgalom, melyben a kutatások együttműködésen alapulnak, átláthatóak és elérhetőek. (openscience.hu)

A nyílt hozzáférésű publikálást mára a legtöbb könyvtár, kutatóintézet és műhely támogatja, a kutatók publikációikat legalábbis részben valamilyen OA konstrukcióban (például Gold, Green, Hybrid) teszik közzé. Az ingyenes hozzáférést azonban valakinek meg kell fizetnie, s a korábbinál nem kevésbé aránytalan helyzetek kialakulásához vezethet, ha ezt maguknak a szerzőknek kell megtenniük. De előidéz sajátos helyzeteket is, különösen a humán tudományok kisebb-nagyobb műhelyeinek folyóiratai esetében. Ezek többsége ugyan semmilyen (elvi, gazdasági) kapcsolatban nincs tényleges OA konstrukciókkal, mégis ezt a kategóriát alkalmazzák magukra, vagy rájuk kiadó intézményük, esetleg szerzőik (pl. impresszumban, intézményi honlapon, Magyar Tudományos Művek Tárában). A kisebb műhelyek maguk bonyolítják a kiadói és szerkesztői tevékenységeket, esetleg intézményi könyvtár segítségét kérik a webes publikáláshoz. Egyetemi és kutatóintézeti könyvtárak például sorra integrálják Open Journal Systems (OJS) platformra a folyóiratokat.¹

Az azonban fel sem merül, hogy ebben a környezetben a szerzőkre terheljék a költségeket, hovatovább ezeknek a folyóiratoknak költségvetésük sincs feltétlenül. A folyóiratszámok munkaköri feladatként, önkéntes munkaként készülnek el, esetleg pályázati finanszírozás teszi azokat lehetővé. Igaz, e tekintetben is egyre szűkül a lehetőségek köre, az NKA Folyóirat-kiadás Kollégiumának akadozó, majd 2016-tól megszűnő tevékenysége sokat rontott a helyzeten. Jelen cikk olvasói számára valószínűleg nem újszerű információk ezek, ám fontos látni, hogy a nagy kiadók és tartalomszolgáltatók, illetve az OA publikálás intézményesült formái között sajátos zárványokat képeznek ezek a kis folyóiratok. Különösen akkor válik mindez érdekessé, amikor a képbe a kutatásértékelés, és a tudásmetria is belép, melyek működtetése szintén javarészt piaci szereplők kezében van.

1 Például ELTE: <http://ojs.elte.hu/>, SZTE: <http://www.ek.szte.hu/ojs/>, MTA KIK: <http://openaccess.mtak.hu/index.php/fejlesztok/ojs>.

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

2018 végén az a – politikai színezetet is kapott – hír rázta meg a tudományos közéletet, hogy az Elektronikus Információszolgáltatás (EISZ) nem hosszabbítja meg a ScienceDirect Freedom Collection, a Scopus és a SciVal adatbázisok hozzáférésére vonatkozó, Elsevierrel kötött szerződését. Ezt az OA politikát támogató szereplők lelkesen, az erről nem tudók, illetve a szkeptikusok a magyar tudományosság gyásznapiaként fogadták, s nem meglepő, hogy utóbbiak médiajelenléte volt intenzívebb.² Az EISZ részletes tájékoztatást közölt az eseményekről:

A jövőben kizárólag az Open Access publikálást 100%-ban támogató, Read & Publish szerződést kívánunk kötni a kiadóval. [...] a magyar kutatói közösség azonban a hozzáférés mellett további előnyökben is részesülne, és mentesülne az OA publikálási költségek alól. (EISZ 2018.)

Lényegében reflektálatlanul maradt később az a hír, hogy az említett adatbázisokhoz való hozzáférést újra lehetővé tették, a kiadóóriással pedig további megállapodásokat irányoztak elő, melyek mind az OA-modell szélesebb körű alkalmazását presszionálják (EISZ 2019a). Ezen felül olyan megállapodások is születtek, melyek az OA publikálás költségei alól mentesítették a szerzőket. (EISZ 2019b és 2019c.)

A kutatási eredmények közzététele azonban nem merül ki tanulmányok közzétételében: azok szerves részét képezik a kutatási adatok is. Bár egyre több bölcsészettudományi projekt vállalkozik (akarva, akaratlanul) valamilyen webes kutatói adatbázis, repozitórium, szövegszolgáltatás elkészítésére, ma is igaz, hogy ezen a területen, illetve a tágabban értelmezett humán tudományokban ez a gyakorlat nem jellemző. A kutatási adatok szabványos megosztását támogatandó az Európai Bizottság nyílt tudománnyal foglalkozó munkacsoportja 2016-ban közzétette a FAIR alapelveket. Az erről készült jelentést és a csoport további dokumentumait szabadon hozzáférhetővé tették (High Level Expert Group on the European Open Science Cloud 2019). Ezen túl számos más kezdeményezés foglalkozik a nyílt tudomány népszerűsítésével és gyakorlatba ültetésével, jelen tanulmányban csak egyet emeltem ki közülük.

A FAIR betűszó a Findable, Accessible, Interoperable és Re-usable szavakból áll össze. Az elvek azt szabályozzák, miként lehet egy (tudományos) dokumentum vagy adat pontosan azonosítható, metaadatokban gazdag, szabványos protokollokon keresztül visszakereshető, hozzáférhető, a tudományos adatok és metaadatok pedig a rendelkezésre álló szótárak/teauruszok segítségével leírtak, valamint újrafelhasználhatók. Az ajánlás csomag lényegében a már működő nyílt hozzáférés és nyílt adat szabványokat, kezdeményezéseket hozza össze, nemcsak javaslatokat fogalmazva meg, de elvárásrendszer is kialakítva ezek köré (Wilkinson és tsai 2016). Az alapelveket például ERC projekt is implementálja.³

De hogy viszonyulnak mindehhez a kutató és a kutatóhelyek? Egyrészt minél szélesebb körű és olcsóbb hozzáférést igényelnek, másrészt azonban féltik saját kutatási adataikat, s gyakran üzleti titokként kezelik az állami vagy EU-s finanszírozásban megszületett eredményeket. Egyrészt széleskörűen alkalmazzák a webes kutatási adatbázisok és tartalomszolgáltatások kínálta digitális szövegeket, másrészt azonban a szerzők töretlenül a papír alapú kiadásokra hivatkoznak – annak ellenére, hogy számos megoldás született már a digitális objektumok azonosítására. Egyrészt még mindig erős (a bölcsészettudományok területén) a digitalizálással kapcsolatos szkepticismus, bizalmatlanság, másrészt azonban a kutatók sokszor kritika nélkül vetik bele magukat a rendelkezésükre álló webes szolgáltatásokba. A tartalom hatalom, s e hatalomért sokfrontos küzdelem zajlik a tudományos szcénán belül is. Ugyanakkor komoly problémát jelent, hogy az egyes adatbázisokban, tartalomszolgáltatásokban kutatva elfelejtjük alkalmazni azt a forráskritikai

² Példaként a Qubit felületén lezajlott vitát említeném: <https://qubit.hu/kereses?q=elsevier>.

³ LiLa: Linking Latin – Knowledge Base of Linguistic Resources (2019) <https://lila-erc.eu/>

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

gyakorlatot, amely az analóg médium esetében már kialakult (Maróthy 2017). A kutatók és kutatóhelyek e kettős természete a tudományos publikálás médiumváltásának, e sokrétű átalakulási folyamatnak a sajátos tünete. A háttérben pedig a tudományos tevékenységet kiszolgáló apparátus (például finanszírozás, közgyűjtemények szolgáltatásai) és az oktatás rendkívül lassú, vagy inkább meg-megszakadó, szigetszerű fejlesztése áll.

Itt megjegyzendő, hogy a webes adatbázisok, tartalomszolgáltatások fenntartói sok esetben elmulasztják jelezni, mi van és mi nincs benne a szolgáltatásban. Gyakori probléma, hogy hiányzik, vagy igen szűkszavú a szolgáltatáshoz tartozó dokumentáció, ez kivált akkor okoz problémát, ha digitalizált anyag közreadásáról van szó, hiszen minden konverzió/médiumváltás adatvesztéssel jár.

Az alábbiakban két példával szeretném megvilágítani, miért is jelent gondot, ha a tudományos eredményekhez csak nyomokban, nem pedig a maguk teljességében férünk hozzá. Egyik példám a weben egyre nagyobb számban elérhető tudományos szövegkiadások esete. A kezdeti HTML-alapú publikálást az ezredforduló után a TEI XML alapú szövegekódolás váltotta fel. A TEI XML a Text Encoding Initiative által gondozott, XML (Extensible Markup Language) szintaxist követő kódkészlet, melyet szövegek kódolására alkalmaznak (TEI: Text Encoding Initiative 2019).

Bár két jelentős vállalkozás is célul tűzte ki a tudományos elektronikus szövegkiadások lajstromba vételét (Franzini és tsai 2018–; Sahle 2008–), ez korántsem egyszerű feladat. Ha a papíralapú szövegkiadások számával vetjük össze, egyik katalógus sem utal számottevő elektronikus kiadványra. Bár a kiadások feldolgozottsága nem teljes, a tudományos elektronikus szövegkiadások számát nagyságrendileg jól mutatja a két katalógus. Az összevetést a papír kiadványokkal az is nehezíti, hogy számos esetben nem egy adott művet tesz közzé a kiadvány, hanem egy akár több száz vagy ezer műből álló korpuszt. Franzini és Sahle gyűjtése tehát nem teljes ugyan, de reprezentatívnak tekinthető. Az alábbiakban Franzini katalógusának adatai alapján tekintem át a forráskódok hozzáférhetővé és felhasználhatóvá tételének gyakorlatát. Egy kutatási projekt keretében ezen adatbázist magyar nyelvű és vonatkozású anyaggal bővítettem, így az elemzés a magyar anyag tekintetében abszolút reprezentatívnak tekinthető (Maróthy 2018b).

A jelenleg 304 tételes katalógus kiadványainak 87,5 %-a szabad hozzáférésű valamilyen szinten, ám csak a teljes anyag 12,5 %-a nyílt forráskódú.⁴ TEI XML alapú kiadványt 169-et számlál (ez a teljes állomány mintegy 55,6%-a), ezek közül 63, tehát közel 37% teszi közzé magukat az XML fájlokat is, itt tehát az arány már biztatóbb. 75 kiadványnak a teljes tartalmára, 3-nak egy részére vonatkozik valamilyen szabad Creative Commons licenc, ez együttesen 46%-a a teljes anyagnak. A forráskódokhoz való szabad hozzáférést támogatják az API-k (Application Programming Interface) is, ilyen felülete 15 kiadványnak van.

A magyar anyagban – mely jelenleg 47 feldolgozott tételt számlál – a kiadványok a szövegeket kivétel nélkül szabadon hozzáférhetővé teszik, ez nagy pozitívum.⁵ GNU General Public License alatt 5 kiadványt találunk, Creative Commons licenc alatt egy sincs. A GNU GPL alattiak kivétel nélkül a *Hálózati kritikai kiadássorozat* (1998–2001) tagjai, melyek egyébként is HTML-ben kódolt szövegek, olyan statikus tartalmak, melyek forrásfájljaihoz egyszerű letöltéssel is hozzáférhetünk. 22 TEI XML kódolású szövegkiadás van, ezek egyike sem tesz közzé semmilyen forráskódot jelenleg, ebben tehát messze elmaradunk a nemzetközi tendenciáktól.

Miért lenne jó, ha az XML-ek hozzáférhetőek lennének? Egyrészt segítségül, mintául szolgálhat más kiadások elkészítéséhez (jelenleg nem tudjuk, hogyan oldanak meg mások ugyanazon

4 Ha a teljes, 304 tételes anyagra vetítjük, a megoszlás a következő: 12,5 % (38) csak fizetés ellenében elérhető, ~3 % (9) regisztrációhoz kötött, de a szöveg ingyenesen hozzáférhető, ~50,6 % (154) a szöveg szabadon hozzáférhető, ~21,4 % (65) a szöveg és a forrásfájl is részben hozzáférhető, 12,5 % (38) a szöveg és a forrásfájl is egészében hozzáférhető. Az adatfelvétel szempontjainak részletes ismertetése: Franzini (2015).

5 Noha a legtöbb magyar kiadvány Franzini gyűjtésében is szerepel, a gyűjtőkör különbségei miatt nem mind (például a már nem elérhetőek hiányoznak), ezért itt saját adatokat használtam (Maróthy 2019). A magyar katalógust utoljára 2019. január 7-én frissítettem.

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

textológiai jelenségeket, például hogyan kódolja egy levél egyes elemeit). Fény derülhetne például arra, hogy bizonyos esetekben a TEI-szabvány nem alkalmas, vagy hiányos (például az egymást átfedő szövegszegmensek miatt új elemet kellett behozni a kódolás során), így közösen járulhatnánk egy elem felvételéért, módosításáért a kódkészletet fenntartó konzorciumhoz. De számos más módon is segíteni tudná a kollaborációt, ha a forrásfájlok minden nehézség nélkül hozzáférhetőek lennének. Másrészt ellenőrizhetővé válna a textológiai, filológiai munka: az XML egyik legfontosabb előnye a kutatási eredmények explicitté tétele a szövegek kódolása során. Ha azonban nem látok rá ezekre az explicitté tett tudományos tényekre, honnan tudjam, mit állít róla, aki rögzítette? Amennyiben az XML fájlokban rögzített adatok csak egy közvetítő grafikus felületen keresztül jutnak el a felhasználóhoz, az jelentősen korlátozza az adatok értelmezhetőségét és felhasználását. Ez a harmadik érvem: a TEI azon alapelve, mely különválasztja a tudományos adatok, a struktúra rögzítését és azok megjelenítését, csak akkor tud érvényesülni, ha a tudományos közösség ezekhez a forrásfájlokhoz hozzá is fér. Az XML kódolás lényege, hogy a tudományos adatokat számítógépes eszközökkel tudjuk kezelni, feldolgozni, elemezni (és megőrizni) – de hogyan tegyük mindezt, ha nem férünk hozzá, mert csak egy olvasókönyvet kapunk?

Másik példám voltaképpen sikertörténet – vagy legalábbis annak kezdete. A számítógépes bölcsészeti kutatások jellemzően nagyobb mennyiségű, ugyanakkor jó minőségű (megbízható és jól strukturált) adatokat igényelnek. Ilyet előállítani időigényes, már rendelkezésre álló adatállományokat megszerezni pedig igen nehézkes. Egy workshop erejéig azonban kísérletet tettünk arra, hogy a különféle webes tartalmak gazdájától adatokat kérjünk, melyeket aztán feldolgozhatunk és elemezhetünk. A téma a Nyugat és korszakának folyóiratai, szerénytelen célunk a folyóiratok és szerzők közötti kapcsolati háló megrajzolása volt (Nyugathon – beszámoló 2019).

A megvalósításhoz a Nyugat és számos más folyóirat szám szintű adataira lett volna szükség, legalább a szerzőnév–folyóiratszám adatelemekre vonatkozóan. Olyan nyilvános adatbázis azonban, amely ezt tartalmazná, jelenleg nem létezik. Így egyfelől a Petőfi Irodalmi Múzeum katalógusát és névterét, valamint az Irodalomtudományi Intézet bibliográfiáját használtuk, másfelől pedig az Arcanum Adatbázis Kft. tudástárát.⁶ Előbbi kettő minőség és strukturáltság szempontjából jónak mondható, mennyiség (pontosabban feldolgozottság) tekintetében azonban kevésnek bizonyult. Utóbbi ugyan nem annyira pontos és strukturált, mint a másik kettő, ám nagyságrendekkel több adatot tartalmaz, így a későbbiekben elsősorban ezzel az adatsorral dolgoztunk. Célszerűbbnek látszott ugyanis egy teljes adatállományt tisztítani, mint egy hiányosat pótolni. Az adattisztítás ebben az esetben azt jelentette, hogy a más funkcióban használt, jobbára következetesen rögzített, de nem strukturált adatokból strukturáltat és feldolgozható készítsünk. Ez egy folyóiratnál hozzávetőleg 2 napot vett igénybe, az adatok pótlása ennél jóval több idő lett volna: a Napkelet esetében például 175 folyóiratszám 991 szerzőjének 4155 cikkéről van szó (Minkó és tsai 2019a és 2019b). E projekt esetében a következő tanulságokat vontuk le: 1) a mai napig számos terület feldolgozottsága alacsony szinten van, így nem képezheti elemzés tárgyát, 2) ha van is megbízható adatállomány, nagyon nehéz hozzájutni, 3) érdemes a már meglévő digitalizált forrásokból kiindulni, azokat javítani még akkor is, ha minőségüket, strukturáltságukat tekintve nem oly bizalomgerjesztők.

A tudományos munka alapfeltétele, hogy a kutatás során feldolgozott adatok visszakereshetők legyenek, az eredmények pedig reprodukálhatók. A digitális bölcsészet esetében például nemrég újra felvetődött a kérdés, vajon az egyes, gyakran kvantitatív elemzéseket tartalmazó publikációk eredményei megismételhetők-e (Da 2019; Underwood 2019). A sokágú vitából itt csupán egy fontos szempont emelnék ki: a tudományos eredmények ellenőrzéséhez és megismételhetőségéhez nélkülözhetetlenek a felhasznált források, fájlok, kódok, parancsok stb. Nem közzétenni például egy

6 PIM Gyűjtemények és Névtér: <https://opac.pim.hu/> és <https://opac-nevtér.pim.hu/>, A magyar irodalomtörténet bibliográfiája: <http://www.iti.mta.hu/mib/>, Arcanum Digitális Tudománytár: <https://adtplus.arcanum.hu/>.

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

adatsort, melyre az elemzés épül, ugyanolyan mulasztás, mint egy felhasznált primer vagy szekunder szakirodalmi tételt nem hivatkozni. Ha a kutatás során felhasznált adatforrás nem áll korábban rendelkezésre, csak akkor láthatjuk, mire alapozza a kutató az állításait, ha ezeket megosztja a tudós közösséggel. Ha nem teszi, tudományos munkája alapjaiban kérdőjeleződik meg. S ez természetesen nem csak a „digitális bölcsészekre” igaz állítás.

A kutatási adatok, adatsorok közzététele mellett tehát számos érv felsorolható, ugyanakkor a jogszabályi / intézményi szabályozási környezet is átalakulóban van. A Magyar Tudományos Akadémia 2012-ben adott ki határozatot a tudományos művek OA közzétételeire (a kutatóintézeteire és a Lendület projektekre vonatkozóan). Mégis feltűnően kevés kutatási adattárral találkozhatunk, aminek egyik oka, hogy az archiválási feladatok jelentős részét az MTA Könyvtár által üzemeltetett repozitóriumok látják el.⁷ Két kivétel azonban van, egyik a 2013-ban indult SZTAKI Publikációtár, másik a Társadalomtudományi Kutatóközpont adatbankja,⁸ mely 2014 óta gyűjti a publikációkat és a kutatási adatokat (Micsik és Gárdos 2014; Meiszterics és Gárdos 2017). Az infrastruktúra és/vagy az egyeztetés hiányát jól szimbolizálja a Zenetudományi Intézetnek az Arcanum Adatbázis Kft. platformjába költözött archívuma, mely nem kapcsolódik az intézményi infrastruktúrákhoz. A fent említett három intézmény a repozitóriumok kezelését tekintve különálló ugyan, de számos ponton kapcsolódik egymáshoz. Mindhárom repozitórium intézményi fenntartású, ugyanazon szoftvert használ (*Eprints*), s mint az idézett Gárdos–Micsik-cikk is mutatja, a kialakításkor együtt is működtek. Ezzel szemben az archívum fejlesztését bemutató Bolya Mátyás (2019) ezt írja:

A szolgáltatáshoz optimalizált fájlméretek és fájlformátumok kellene, valamint megfelelő – vagyis fenntartható, fejleszthető, korlátlan elérésű – szoftverkörnyezet. Ehhez kapcsolódik a megfelelő publikációs platform – és üzemeltetési konstrukció – kialakítása. A tapasztalatok szerint egy ilyen platform üzemeltetése meghaladja egy kutatóintézet lehetőségeit, ezért vezető piaci szereplőket kell bevonni a munkába.

A kutatóintézeteknek a szabályozás szerint kötelezettsége a tudományos műveket OA közzétételi rendszerben kezelni, megőrizni és hozzáférhetővé tenni. Az ehhez szükséges infrastruktúrát azonban az intézmény nem határozza meg pontosan, ahogyan azt sem, mi tartozik a tudományos művek közé. (Ez az intézmény eddig a Magyar Tudományos Akadémia volt – a helyzetet tovább bonyolítja, hogy már nem az, vagy csak részben az.) A 2012-es határozat tárgyi hatálya:

jelen határozat az (1) bekezdésben meghatározott kutatók által munkaköri kötelezettségként a tudományos tevékenységükkel összefüggésben alkotott, a Magyar Tudományos Művek Tára (MTMT) alapján tudományos műnek minősülő közlemények kezelésére, megőrzésére és hozzáférhetővé tételére terjed ki. (Irányelvek 2012: 1.)

A 2016-os határozat tárgyi hatálya ehhez képest így módosult:

jelen határozat az (1) bekezdésben meghatározott kutatók által közpénzekből finanszírozott és munkaköri kötelezettségként, a tudományos tevékenységükkel összefüggésben alkotott, a szerzői jogról szóló mindenkor hatályos szabályok szerint tudományos műnek minősülő alkotások (a továbbiakban: mű) kezelésére, megőrzésére és hozzáférhetővé tételére terjed ki. (Irányelvek 2016: 231.)

7 REAL – az MTA Könyvtárának Repozitóriuma: <http://real.mtak.hu/> ; REAL-J – az MTA Könyvtárának Repozitóriuma: <http://real-j.mtak.hu/>

8 SZTAKI Publikációtár (2013–): <http://eprints.sztaki.hu/>; MTA TK Kutatási Dokumentációs Központ (2014–): <https://www.kdk.tk.mta.hu/>

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

A módosítás egyrészt beemeli a tudományos művek közé a közpénzekből finanszírozott munkákat (vagy az állami finanszírozás tényét nyomatékosítja), másrészt megváltoztatja a tudományos mű meghatározását. Míg a korábbi változatban az MTMT, a későbbiben a szerzői jog az irányadó, mely a tárgyi hatály jelentős kiterjesztését mutatja. Vagy mégsem? A szerzői jog konkrétan nem a „tudományos mű”-ről, hanem a tudományos alkotásokról rendelkezik, bár a „mű” és „alkotás” fogalmakat lényegében szinonimaként használja. Kérdés, az MTA-határozat „tudományos mű”-fogalma csak a publikációkra értendő (a nyilvánosan, valamilyen analóg vagy digitális platformon közzétett művekre), vagy egyenértékű a szerzői jogi törvény tágabb, „tudományos alkotás” fogalmával. Utóbbi esetben az MTA-határozatnak nemcsak a publikációkra kell vonatkoznia, hanem például ezekre is: „c) a számítógépi programalkotás és a hozzá tartozó dokumentáció (a továbbiakban: szoftver) akár forráskódban, akár tárgykódban vagy bármilyen más formában rögzített minden fajtája, ideértve a felhasználói programot és az operációs rendszert is”, illetve „p)⁹ a gyűjteményes műnek minősülő adatbázis” („1999. évi LXXVI. törvény a szerzői jogról” I. fejezet). De mi az az adatbázis?

60/A. § (1) E törvény alkalmazásában adatbázis: önálló művek, adatok vagy egyéb tartalmi elemek valamely rendszer vagy módszer szerint elrendezett gyűjteménye, amelynek tartalmi elemeihez – számítástechnikai eszközökkel vagy bármely más módon – egyedileg hozzá lehet férni. (1999. évi LXXVI. törvény a szerzői jogról: VII. fejezet.)

Ennek megfelelően mindegy, hogy egy Excel-táblába gyűjtött prozopográfiai adatsorról (személyi adattárról), vagy egy egyéni fejlesztésű tartalomkezelő rendszerben weben publikált multimédiás kutatási gyűjteményről van szó, az a szerzői jog hatálya alá tartozik. Ha tehát kutatóintézeti dolgozóként vagy Lendület-pályázat kutatójaként egy majdani szaktanulmányom számára szoftveres támogatással rendszerezett adatokat gyűjtök, vagy webes tudományos szövegkiadást készítek, köteles vagyok/voltam (hogy melyik, az a jelenlegi jogi környezetben mindenki számára talány) az adatsort forrásfájlokkal együtt az MTA határozatnak megfelelően OA-modellben közzétenni. Mert közpénzből készült, a köz számára. Nem árt hangsúlyozni, hogy a tudományos műveket nem az MTA-nak kell átadnom, hanem szabadon hozzáférhetővé kell tennem a közösség számára. A határozat a tekintetben rugalmas, hogy ezt a kutató miképpen valósítsa meg. Ehhez készült egy segédlet is az MTA Könyvtárában, mely valamelyest behatárolja a tudományos művek körét. E szerint csak teljes tudományos művekről van szó, illetve változatlanul az MTMT az irányadó (Bilicsi 2017).

Ennyit a jogi környezetről. A fenti példákból is látszik azonban, hogy a gyakorlat valami mást mutat: a sorra készülő, kutatási projektekhez kapcsolódó digitális alkotásokat részint rejtegetik, részint csak bizonyos vetületükhöz kínálnak hozzáférést. A humán területeken folyó digitalizálási, adatbázis-építési munkálatok egyébként is roppant széttagoltságot mutatnak, nehéz kideríteni, mi készült el (akár többször is), és mi az, aminek a feldolgozása még várat magára. A szigetszerűen működő tudományos adatbázisok, szövegszolgáltatások fenntarthatóságával, megőrzésével pedig csak a legkritikább esetben foglalkoznak az erre illetékes szakemberek. A megőrzést, megőrződést is segítheti a szabad hozzáférés biztosítása, ezt a webarchiválási projektek jól bizonyítják (Drótos 2017; Kokas és Drótos 2018). Az első hazai hálózati tudományos szövegkiadást például már csak az Internet Archive-ból tudtam visszaállítani – a hálózati Balassi-kiadások egyébként is kalandos történetet tudhatnak maguk mögött (Tóth 2004; Maróthy 2018a). Ezeket a problémákat valóban jobban tudja kezelni egy megfelelő intézményi repozitórium, mely nemcsak PDF-ben feltöltött

⁹ Ez a kiegészítés 2002. január 1-jével lépett hatályba. Az adatbázisra vonatkozó részletes rendelkezéseket a VII. fejezet tartalmazza.

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

szakcikket, hanem kutatási adatokat, multimédiás anyagokat tartalmazó különféle formátumokat is képes kezelni, és strukturált, könnyen hozzáférhető, megtalálható, felhasználható módon közzétenni a weben. Hangsúlyozandó, hogy a legkevésbé sem a centralizálás (egy mindent magába foglaló hipotetikus adatbázis) a megoldás a problémákra, hanem az említett nemzetközi szabványok alkalmazása a helyi igényeknek megfelelően, valamint az egyes érdekközösségek (intézmények, csoportok) összefogása, együttműködése.

De miért nem teszi közzé a többség a kutatási adatokat és a forrásfájlokat? Túl azon, hogy döntéseinket számos esetben a tudatlanság, illetve az információhiány befolyásolja, azért, mert félünk. Hogy valaki elveszi az adatot és tönkretesz, hogy elhalássza az ötletet, vagy jobbat talál ki helyette, vagy jobban csinálja meg, stb. Mert megfoszt a dicsőségtől, s még inkább a hatalomtól: hogy ez a tartalom csak az enyém, következésképpen az előnyeiket is csak én élvezhetem.

Sok esetben természetesen nem ilyen egyszerű a helyzet: például egy ötéves alap kutatás után valamelyest érhető, ha az elemzés elsőbbségét valaki magának vagy a kutatócsoportjának tartja meg. Az esetek és az elvárások sokfélék. Az aggodalmak nagy részét azonban eloszlatja, ha az OA-modellen alapuló kutatási repozitórium közzétételi lehetőségeiről, illetve a rendelkezésre álló licencekről (például az említett GNU-ról és CC-ről) tájékozódunk. Bizonyos alkotásokat késleltetve is közzétehetünk, indokolt esetben beiktathatunk személyes adatigénylést, s ami a legfontosabb: a felhasználási feltételeket korlátozhatjuk – például tilthatjuk a módosítást, vagy a kereskedelmi célú felhasználást. Érdemes azonban minden korlátozás érvényesítése előtt átgondolni, hogy valóban indokolt-e, illetve előnyös-e egyáltalán akár az egyén, akár a (tudományos) közösség számára. Minél több a tiltás, annál nagyobb az esélye, hogy az eredmények nem hasznosulnak, azokat nem használják, nem hivatkoznak rájuk, vagy nem is hallanak róla. Elvész a kollaboráció megannyi lehetősége, s amíg az eredmény nem ellenőrizhető, nincs visszajelzés a kutatás sikeressége, minősége tekintetében. Ez összességében nagyban gátolja a tudomány fejlődését.

Ha nem is az a cél, hogy mindenki mindenkivel mindent megosszon bármilyen felhasználásra, vitán felül áll, hogy a kutatási eredményekhez (kivált ha közpénzből finanszírozták azokat) széles körű hozzáférést kell biztosítani. Szerencsére ehhez a technikai feltételek jobbra már adták, így nincs más teendő, mint átgondolni és megbeszélni, mit és hogyan osszuk meg (például adattárolási és -csere szabványok, licencek), kivel osszuk meg ezeket (szélesebb körben vagy csak a kutatói közösséggel), s milyen előnyökkel és hátrányokkal jár ez a nyitottság.

Irodalom

1999. évi LXXVI. törvény a szerzői jogról. Hatályos Jogszabályok Gyűjteménye (2019): <https://net.jogtar.hu/jogszabaly?docid=99900076.TV>
- Irányelvek 2012 = A Magyar Tudományos Akadémia elnökének 27/2012. (IX. 24.) számú határozata: Tudományos művek nyílt hozzáférésű közzétételének irányelvei (2012): https://mta.hu/data/dokumentumok/hatteranyagok/akademiai_szabalyozasok/27_2012_elnoki_hat_Open_Access5.pdf
- Irányelvek 2016 = A Magyar Tudományos Akadémia elnökének 24/2016. (VII. 15.) számú határozata: A tudományos művek nyílt hozzáférésű közzétételéről (2016): http://openaccess.mtak.hu/dokumentumok/mta_határozat_oa_2016.pdf
- Bilicsi Erika (2017) Az MTA open access határozatában előírtak teljesítése. http://openaccess.mtak.hu/dokumentumok/mta_oa_hat/mta_oa_bilicsi.pdf
- Bolya Mátyás (2019) Interaktív térképen a magyar népzene – úton a digitális bölcsészet felé. *MTA.hu*, 2019. május 20. https://mta.hu/tudomany_hirei/interaktiv-terkepen-a-magyar-nepzene-uton-a-digitalis-bolcseszet-fele-109688

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

- Da, Nan Z. (2019) The Computational Case against Computational Literary Studies. *Critical Inquiry*, 45(3): 601–639. <https://doi.org/10.1086/702594>
- Drótos László (2017) Az internet archiválása mint könyvtári feladat. *Tudományos és Műszaki Tájékoztatás*, 64(7–8): 361–71.
- EISZ (2018) Kérdések és válaszok az Elsevier tárgyalásokról. <http://eisz.mtak.hu/index.php/hu/open-access/281-kerdesek-es-valaszok-az-elsevier-targyalasokrol.html>
- EISZ (2019a) Szándéknyilatkozat aláírásával indultak újra a tárgyalások az EISZ és az Elsevier között. <http://eisz.mtak.hu/index.php/hu/open-access/337-szandeknyilatkozat-alairasaval-indultak-ujra-a-targyalasok-az-eisz-es-az-elsevier-kozott.html>
- EISZ (2019b) Open access megállapodás az Akadémiai Kiadóval. <http://eisz.mtak.hu/index.php/hu/open-access/295-open-access-megallapodas-az-akademiai-kiadoval.html>
- EISZ (2019c) Read & Publish megállapodás a Royal Society of Chemistry-vel. <http://eisz.mtak.hu/index.php/hu/open-access/313-read-publish-megallapodas-a-royal-society-of-chemistry-vel.html>
- Franzini, Greta (2015) How to contribute a digital edition to the Catalogue. 2019. június 3. https://github.com/gfranzini/digEds_cat
- Franzini, Greta és tsai (2018–) Catalogue of Digital Editions. <https://dig-ed-cat.acdh.oeaw.ac.at/>
- High Level Expert Group on the European Open Science Cloud (2019) Open Science – Research and Innovation – European Commission. <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud-hleg>
- Kokas Károly és Drótos László (2018) Webarchiválás és a történeti kutatások. *Digitális Bölcsészet*, 1: 35–54. <https://doi.org/10.31400/dh-hun.2018.1.129>
- Maróthy Szilvia (2017) Elektronikus forráskritika. *Könyv, könyvtár, könyvtáros*, 26(6): 15–20.
- Maróthy Szilvia (2018a) A tudományos elektronikus szövegkiadások feldolgozása. In Kiszl Péter és Csík Tibor, szerk.: *Valóságos könyvtár – könyvtári valóság III*. Budapest, ELTE BTK Könyvtár-és Információtudományi Intézet. 351–356.
- Maróthy Szilvia (2018b) Tudományos szövegkiadások a hálózaton: Áttekintés. *Irodalomtörténeti Közlemények*, 122(5): 617–633.
- Maróthy Szilvia (2019) Tudományos elektronikus szövegkiadások bibliográfiája. *GitHub*. https://github.com/marothyszilvi/digEds_hun
- Minkó Mihály és tsai (2019a) A Nyugat és korszakának teljes hálózata helyett (Kapcsolati háló felrajzolásának nehézségei). Előadás a *Hálózatok és komparatiztika: a világirodalom hálózatai* c. konferencián, Kolozsvár, 2019. október 4.
- Minkó Mihály és tsai (2019b) Nyugat. *GitHub*. <https://github.com/marothyszilvi/nyugat>
- Meiszterics Enikő, és Gárdos Judit (2017) (Át)láthatóbb kutatások – az MTA TK Kutatási Dokumentációs Központ tapasztalatai. *Tudományos és Műszaki Tájékoztatás*, 64(7–8): 389–96.
- Micsik András, és Gárdos Judit (2014) Tudományos repozitóriumok az MTA-ban: a KDK és a SZTAKI tanulságai. In Kunkli Roland, Papp Ildikó, Rutkovszky Edéné, szerk., *Informatika a felsőoktatásban*. Debreceni Egyetem Informatikai Kar. 773–781. <http://openarchive.tk.mta.hu/340/>
- MTA Open Access rendelkezés (é. n.): <http://real.mtak.hu/eprints/mandate.html>
- Nyugathon – beszámoló (2019) *Digitális Bölcsészet weboldala*, 2019. március 22. <http://ojs.elte.hu/index.php/digitalisbolcsesz/announcement/view/15>
- openscience.hu = *Mi az Open Science?* <https://openscience.hu/mi-az-open-science/>
- Sahle, Patrick (2008–) A catalog of Digital Scholarly Editions, v 3.0. <http://www.digitale-edition.de/>
- TEI: Text Encoding Initiative (2019): <https://tei-c.org/>

PREPRINT

Maróthy Szilvia, „A nyílt és a zárt tudományról”, In *Kulturális iparágak, kánonok és filterbuborékok*, szerk. Bárány Tibor, Hermann Veronika, Hamp Gábor, Typotex kiadó, 2020, megjelenés alatt.

Tóth Tünde (2004) Online kritikai szövegkiadás Magyarországon az ezredfordulón. *Helikon*, 3: 417–40.

Underwood, Ted (2019) Dear Humanists: Fear Not the Digital Revolution. *The Chronicle of Higher Education*, 2019. március 27.
<https://www.chronicle.com/article/Dear-Humanists-Fear-Not-the/245987>

Wilkinson, Mark D. és tsai (2016) The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3: 160018. <https://doi.org/10.1038/sdata.2016.18>