

Holl András

REAL – az MTA Könyvtárának repozitórium

Az MTA Könyvtára széles körű elektronikus könyvtári szolgáltatásokat nyújt – ezek között jelent meg 2009-ben egy repozitórium, a REAL, amely jelenleg az OTKA által átadott, 2006–2009 között elfogadott kutatási jelentéseket tartalmazza.

Az e-könyvtári szolgáltatások között szereplő elektronikus gyűjtemények és a repozitóriumok sok szempontból a szolgáltatási spektrum átellenes oldalán találhatók. Az előbbiek, mint például az *MTA Könyvtárának (MTAK)* digitális gyűjteményei, a *Carbo kódex*¹, a *Kőrösi Csoma*², a *Bolyai*³ vagy a *Radnóti*⁴ anyag egyedi, különösen értékes dokumentumokat tárnak a nagyközönség elé, igényes grafikai kivitelezésben, gondos feltáró munka nyomán. Egy repozitórium anyagai inkább a tudóstársadalomnak szólnak, mint a laikus közönségnek, nagy mennyiségű dokumentumot tartalmaznak teljes szöveggel, és inkább adatbázisként, mint kiállításként működnek.

Az MTAK elektronikus könyvtári szolgáltatásai között 2009-ben jelent meg egy új repozitórium, a *REAL* (az MTAK repozitórium = **RE**pository of the **Academy's** Library).

A repozitórium létesítésének indítékai között az elektronikus könyvtári szolgáltatások, ezen belül a teljes szövegű anyagok bővítésének szándéka, és a hagyományos gyűjtemények, mint például a

doktori dolgozatok elektronikus változata megteremtésének igénye szerepelt.

A REAL létrejötte

Az OTKA (*Országos Tudományos Kutatási Alapprogramok*) 2009-ben olyan szabályozást fogadott el⁵, miszerint minden általa támogatott kutatóprogram tudományos eredményeit a *Nyílt Hozzáférés (Open Access = OA)* szabályai szerint hozzáférhetővé kell tenni. Azt ugyan nem szabta meg, hogy ezt milyen módon szükséges megtenni – valamely repozitóriumban való elhelyezéssel, OA folyóiratban, vagy hagyományos folyóiratban az OA lehetőség megvásárlásával –, de biztosítani kívánt egy olyan repozitóriumot, amelyben a cikkek – más lehetőség híján – mindenképpen elhelyezhetők. Ezen repozitórium felállítására kérték fel az MTA Könyvtárát (1. ábra). A szolgáltatás kiépítését anyagilag is támogatták. (Ugyanezt a repozitóriumot választotta az OTKA az általa támogatott kutatóprogramok pályázati zárójelentéseinek nyilvánosságra hozatalára és megőrzésére is.)



1. ábra A REAL nyitóoldala

A létesítendő szolgáltatásnak az *Open Archives Initiative* (OAI = *Nyílt Hozzáférés Kezdeményezés*) kritériumainak megfelelően kellett felépülnie, és alapvető elvárás volt vele szemben a nemzetközi rendszerekhez való kapcsolódás képessége is.

A rendelkezésre álló támogatás mértéke és a könyvtár adottságai alapján, az MTAK szabad szoftver alkalmazása mellett döntött. A számos rendelkezésre álló szoftver közül az *EPrints*, *DSpace*, *CDS Invenio* és a *Fedora Commons* rendszereket vizsgáltuk meg, s végül egyszerűsége miatt az elsőt választottuk. EPrints-alapú repozitóriumok már működtek az országban, igaz, a szoftver korábbi verziójával, így hazai segítségre is támaszkodhattunk. Saját tapasztalatainkat, az új változat magyar fordítását ugyanakkor mi is igyekszünk az érdeklődők rendelkezésére bocsátani. Ennek megfelelő fórumot teremt a *HUNOR* (*HUNGarian Open Access Repositories*) egyesület⁶ és az NIIF különböző rendezvényei: az *IPSZILON szeminárium*⁷ és a *NETWORKSHOP*⁸.

A szolgáltatáshoz dedikált szerverszámítógépet szereztünk be, sőt, még egy tartalék gépet is beállítottunk, valamint kiépítettünk egy más feladatra is felhasználható, bővíthető, nagy megbízhatóságú, SAN hálózaton keresztül elérhető tárolórendszert. A szerver Linux operációs rendszer alatt működik.

Az OTKA-gyűjtemény

A repozitórium építése 2008-ban kezdődött, maga a szolgáltatás pedig 2009 nyarán indult meg. Az új repozitórium neve egy angol betűszó lett: REAL (Repository of the Academy's Library).

A REAL, megnyitása idején, körülbelül kétezer dokumentumot tartalmazott: az utóbbi években lezárult OTKA-projektek zárójelentéseit, amelyeket az OTKA Iroda helyezett el – és helyezi majd el a jövőben is – a REAL-ban. Terveink szerint a feltöltés automatikus lesz majd, vagyis az OTKA Elektronikus Pályázati Rendszeréből a SWORD protokoll segítségével kerülhetnek majd át a jelentések a repozitóriumba.

Tudományos folyóiratcikkek még csak nagyon kis számban találhatóak a REAL OTKA-gyűjteményében. Ennek fő oka, hogy az OTKA OA szabályozásának életbe lépése után elfogadott projektek még nemigen eredményeztek cikkeket. A szakcikkeket maguknak a szerzőknek kell feltölteniük a REAL-

ba, amit annak webes felületén keresztül tehetnek majd meg. Ehhez először azonosítót kell készíteniük maguknak (hasonlóan ahhoz, ahogyan egy levelezőrendszerbe vagy közösségi oldalra regisztrálnának). Az azonosító birtokában a fájlok feltöltése, a metaadatok megadása nagyjából tíz perces munkát jelent. Az EPrints szoftver itt is kínál egyszerűt: a DOI (*Digital Object Identifier*) import funkcióval⁹ csak ezt az egyedi azonosítót kell megadni, a többi metaadatot a szoftver lekérdezi a *CrossRef*¹⁰ adatbázisából. A könyvtár munkatársai, természetesen, szerzői feltöltés esetén is ellenőrzik, hogy a tétel a repozitórium gyűjtőkörébe tartozik-e, megfelelőek-e a metaadatai, és probléma vagy kérdés esetén segítséget nyújtanak a szerzőknek.

Az EPrints szoftver keresési lehetőségeket is kínál: egyszerű vagy összetett kereséssel. Ez utóbbival címre, szerzőre, a tartalmi kivonat szavaira, szakterületre, tételtípusra (cikk, könyv fejezete, dokumentáció stb.), állapotra (megjelent, in print, beküldve, kézirat), folyóirat vagy kiadvány címére stb. egyaránt kereshetünk.

A REAL-ban tárolt tételeket azonban az *OAI-PMH protokoll* (*Open Access Initiative. Protocol for Metadata Harvesting*)¹¹ alkalmazása miatt valószínűleg más keresők, aggregátorok segítségével fogják majd megtalálni az érdeklődők (jelenleg is kereshető például az *nda@sztaki* segítségével).

A REAL háttértárat nyújt más projekteknek is, mint például az OTKA-nak vagy az *MTMT-nek* (*Magyar Tudományos Művek Tára*)¹², és természetesen a repozitóriumok nemzetközi lajstromaiba (mint pl. a *ROAR* (*Registry of Open Access Repositories*)¹³, *OpenDOAR* (*The Directory of Open Access Repositories*)¹⁴ is bejegyezték.

A REAL elérhető az MTAK weboldaláról, és az OTKA OA anyagai is hivatkoznak címére¹⁵.

A disszertációgyűjtemény

Az OTKA-gyűjtemény után (amely, mint láthattuk, az OTKA-projektek zárójelentéseit és tudományos eredményeit, többnyire folyóiratokban megjelenő cikkeket tartalmaz) az *akadémiai doktorok disszertációit tartalmazó gyűjtemény* nyílik majd meg. (Ebben helyeznek el egyes kandidátusi dolgozatokat is.)

Ennek a gyűjteménykörnek az ötletét már 2003-ban felvetette *Domsa Károlyné*, az MTA Könyvtára

akkori vezetője (Bánhegyi, 2003). A doktori gyűjtemény neve REAL-d lett, s külön címen érhető el az MTAK honlapján¹⁶.

A gyűjtemény feltöltése ez esetben is több csatornán keresztül történik:

- A nyomtatott, gépelt dolgozatok teljes digitalizálására nem gondolhatunk – ez a jelenleg rendelkezésre álló lehetőségekkel nem valósítható meg, de valószínűleg igény se lenne rá. A gyakorta gyenge minőségű gépelés, papír, kötés miatt a disszertációk digitalizálása nagy kihívást jelent, különösen úgy, hogy az eredeti könyvet is meg kell őrizni, felválni nem lehet. Ezért az igény szerinti (on demand) digitalizálás stratégiáját választottuk: azokat az értekezéseket szkenneljük, amelyeket olvasóink kerestek és keresnek. (A munkához olcsó, nem automatizált könyvszkennert használunk.)
- A jövőben fénymásolási igényeket már nem teljesítünk a hagyományos módon: az igényelt művet szkenneljük és kinyomtatjuk – és egyúttal PDF formában a repozitóriumban is elhelyezzük. A szerző maga is kérheti doktori vagy kandidátusi dolgozatának digitalizálását – és be nem kötött, vagy felvágható példány beküldésével segítheti is azt! Ily módon évente nagyjából száz dolgozat digitalizálása történhet meg.
- Az elmúlt két évtizedben keletkezett dolgozatok esetében még van esély arra, hogy a fájlok fellelhetők legyenek. Az első lépések az írások összegyűjtésére reménykeltők: már majd' félszáz dolgozatot sikerült beszerezni. Az elektronikus levélben vagy optikai lemezen beküldött dolgozatokat a könyvtár munkatársai töltik fel a repozitóriumba, bár természetesen itt is van mód arra, hogy a kutató saját azonosítót készítsen, és maga töltsse fel értekezését. Fontos a gyors cselekvés: az idő múltával az elektronikus változatok beszerzésének lehetősége csökken.
- A jövőben az újonnan elfogadott dolgozatok feltöltését a *Doktori Tanács Titkársága* végezheti majd – eleinte manuálisan, később elektronikus rendszerük ugyancsak a SWORD felületen (l. Holl et al., 2009) keresztül tölthet majd fel anyagokat a REAL-d-be.
- A dolgozatok mellett a REAL-d ún. járulékos dokumentumokat is tárol: téziszüzeteket, bírálatokot és azokra adott válaszokat is. Ez utóbbiak azonban nem feltétlenül olvashatók nyilvánosan.

Alapelvek, tanulságok

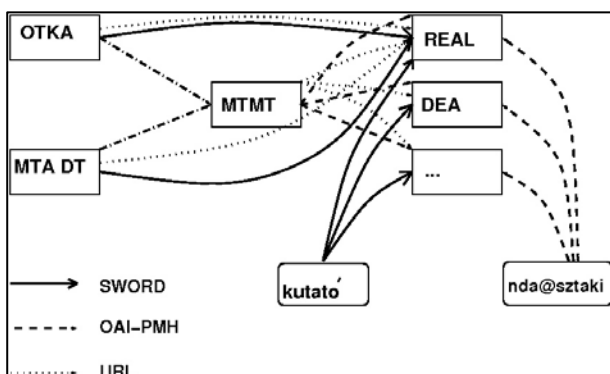
A repozitórium létrehozásakor lefektettünk néhány alapelvet:

- Első gyűjteményeink tudományos jellegűek, melyekben valamilyen módon ellenőrzött minőségű anyagokat kívánunk tárolni: az első gyűjtemény esetében (REAL) az OTKA Iroda által elfogadott kutatási jelentéseket vagy folyóiratokban közölt (bár nem feltétlenül referált) cikkeket, a másodikonál (REAL-d) elnyert tudományos fokozatokhoz tartozó értekezéseket. A könyvtárnak azonban nincs módja az elhelyezett anyagok minőségének megállapítására, így abban külső szűrőkre hagyatkozunk.
- Nem lenne jó azonban csak szórványos feltöltöttségű repozitóriumok kialakítása. Az OTKA-támogatású kutatóprogramok folyóiratcikkeinek ugyan csak egy része fog a REAL-ba kerülni, ám itt is törekszünk arra, hogy az nagy mennyiségű cikket tartalmazzon. A doktori dolgozatok és OTKA-jelentések esetében viszont már azt a célt is kitűzhetjük, hogy – egy adott dátumtól kezdve – teljes legyen a gyűjtemény.
- Az első tapasztalatok azt mutatják, hogy nagy gondot kell fordítanunk az elhelyezett dokumentumok minőségbiztosítására, legalábbis arra a részére, amelyben a könyvtárnak kompetenciája van: a megfelelő metaadatokra, a hitelességre, a formátumra. Alapvető fontosságú, hogy a szükséges metaadatok minden esetben rendelkezésre álljanak, s minél kevesebb hibát tartalmazzanak. Fontos továbbá az esetlegesen meglévő egyedi azonosítók (mint pl. a DOI) tárolása is. A metaadatoknak nemcsak a repozitórium mögött álló adatbázisban kell szerepelniük, hanem a PDF állományok metaadat-mezőiben, valamint maguknak a dokumentumoknak a szövegében is, hiszen a cikkek, dolgozatok a repozitóriumkontextusból kiszakíthatók: letölthetők és kinyomtathatók, s a fontosabb metaadatoknak akkor is folyamatosan kísérniük kell a dokumentumokat. (Felmerült valamiféle integritás-ellenőrző megoldások használata is.)
- A hosszú távú megőrzés szempontjából fontos az elfogadott állományformátumok megválasztása is. A sokféle formátum megkönnyíti a feltöltők dolgát, ám idővel egyre nagyobb terheket ró a könyvtárra. Éppen ezért az MTAK csak a PDF formátumot engedi meg, annak is az archiválásra tervezett PDF/A¹⁷ változatát részesíti előny-

ben. A szolgáltatás és a megőrzés érdekei, persze, néhol ütköznek: az utóbbi szempontjából az egyszerűség és szabványosság a kulcsfontosságú.

Szolgáltatások hálójája

A REAL nem magában álló szolgáltatás – az OAI-PMH protokollt alkalmazó repozitóriumokat eleve más szolgáltatásokkal való összekötésre tervezték (2. ábra). De ez nem csak az arathatóságról szól. A hazai tudományos – felsőoktatási – közgyűjteményi kör elektronikus szolgáltatásait oly módon kell hálózatba szervezni, hogy minél magasabb színvonalú, minél gazdagabb szolgáltatást nyújtsanak; a szolgáltatás minél költséghatékonyabb legyen, és mindez a kutatók, intézmények számára minél kisebb terhet jelentsen.



2. ábra Szolgáltatások hálójata

A repozitóriumok tárolhatják és szolgáltatathatják a dokumentumokat, a bibliográfiai adatbázisok – mint például az MTMT – a szcientometriai adatokat (publikációs listákat és a publikációkra kapott idézeteket), a minősítési és bírálati eljárások során pedig a teljes szövegű dokumentumokra hivatkozó publikációs listák is felhasználhatók. Az aggregátorok és keresők – mint az *nda@sztaki* – arathatják a repozitóriumok metaadatait, keresési lehetőségeket nyújthatnak a felhasználóknak. A specializáció mellett ugyanakkor redundanciák megteremtésére is szükség lehet: az NIIF például háttértárat vagy biztonsági mentést tesz lehetővé ezeknek a szolgáltatásoknak (l. *Stefán Péter* előadását ugyanezen a konferencián, 2010. március 17-én).

A szolgáltatásháló elemeinek összekapcsolását különböző szabványok alkalmazása segítheti: például a már említett OAI-PMH és SWORD, a DublinCore, az OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*)¹⁸.

Természetes, hogy a heterogén, elosztott rendszereknél nehézségek is adódhatnak: nagyobb problémát jelenthet például a duplumszűrés (ehhez adhat segítséget az egyedi azonosítók használata).

A repozitóriumok dokumentumokkal való feltöltése, maga a tartalomszolgáltatás fontos kérdés. A REAL esetében, szerencsére, bizonyos dokumentumtípusoknál – OTKA zárójelentések, doktori értekezések – megoldott a folyamatos gyarapodás. A szakcikkek esetében már más a helyzet, itt még a Nyílt Hozzáférés OTKA által való megkövetelése sem jelent biztosítékot: nem minden cikk kerül majd a REAL-ba.

Fontos feladat tehát a kutatók *elhelyezési hajlandóságának* növelése. Ez történhet technikai úton: mind a már említett, jelenleg is rendelkezésre álló DOI import funkció ajánlásával, mind pedig a SWORD alapú, egyszerű feltöltést lehetővé tevő eszközök kifejlesztésével. De talán ezeknél is fontosabb a tájékoztatás (Holl 2010).

Sok kutató nincs tisztában azzal, hogy az OA már szinte minden kiadó politikájával összeegyeztethető: nem kell publikálási szokásainkon változtatni, s az OA közzététel sem kell, hogy jogsértéssel vagy többletkiadással járjon. A dokumentum elhelyezése előtt a REAL is ajánlja a *SHERPA RoMEO*¹⁹ szolgáltatást a kiadói politika ellenőrzésére.

A REAL – az EPrints által nyújtott funkcióval – nemcsak embargóperiódus beállítását teszi lehetővé a kiadói érdekek védelmére, de zseniális technikai lehetőséget is ad a jog által biztosított lehetőségek kiaknázására: ha a feltöltő szerző megadta elektronikus levélcímét – amely természetesen a felhasználók számára nem látható –, egy gombnyomással az embargós cikkekből is lehet a szerzőtől másolatot kérni. A szerzőnek nincs egyéb dolga, mint hogy a repozitórium által küldött levélben az engedélyező hiperhivatkozásra kattintson.

Megfelelő szolgáltatáshálózatot kiépítve a szerzőnek egy megfelelő repozitóriumban elhelyezett cikkével már nem kell többet foglalkoznia: az automatikusan bekerül publikációs listájába, és tudományos fokozatszerzésnél, habilitációnál, pályázatbeadásnál vagy jelentéskészítésnél a befektetett munka megtérül.

Hivatkozások

- ¹ Ludovicus Carbo: De divi Mathiae regis laudibus rebusque gestis dialogus: <http://carbo.mtak.hu>
- ² Kőrösi Csoma Sándor, a tibetisztika megalapítója és hagyatéka az MTAK honlapján: <http://csoma.mtak.hu>
- ³ A Magyar Tudományos Akadémia Könyvtárának Bolyai János honlapja: <http://bolyai.mtak.hu>
- ⁴ Száz éve született Radnóti Miklós. Digitális kiállítás az MTAK honlapján: <http://radnoti.mtak.hu>
- ⁵ http://www.otka.hu/index.php?akt_menu=106&hir_reszlet=133
- ⁶ <http://www.open-access.hu/index.php/HUNOR>
- ⁷ <http://ipszilon.niif.hu>
- ⁸ http://www.niif.hu/hu/networkshop/niif_networkshop_konferenciak
- ⁹ <http://www.eprints.org/software/training/users/depositing.php>
- ¹⁰ <http://www.crossref.org/>
- ¹¹ <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- ¹² <http://www.mtmt.hu/>
- ¹³ <http://roar.eprints.org/>
- ¹⁴ <http://www.opendoar.org/>
- ¹⁵ <http://real.mtak.hu>
- ¹⁶ <http://real-d.mtak.hu>

¹⁷ <http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>

¹⁸ <http://www.openarchives.org/ore/>

¹⁹ <http://www.sherpa.ac.uk/romeo/>

Irodalom

BÁNHEGYI Zs.: Nyílt Hozzáférés Kezdeményezés (Open Access Initiative) – Kitekintés és körkép. = TMT, 50. köt. 6–7. sz. 2003. p. 236–249.

HOLL, A. – KARÁCSONY Gy. – HARANGI G. – CSERBÁK M.: Teljes szövegű publikációk cseréje repozitóriumok között – a SWORD protokoll tesztelése az MTA Könyvtárban és a Debreceni Egyetem Könyvtárban. = NETWORKSHOP 2009. <https://nws.niif.hu/ncd2009/docs/ehu/035.pdf>

HOLL A.: Nyílt hozzáférés a tudományos szakirodalomhoz – hazai fejlemények. = Magyar Tudomány, 171. köt. 1. sz. 2010. p. 58–61.

Beérkezett: 2010. IV. 19-én.



Holl András

az MTA Könyvtára informatikai tanácsadója,
az MTA Konkoly Thege Miklós Csillagászati Kutatóintézetének munkatársa.
E-mail: holland@mtak.hu

Egyre több időt töltünk a közösségi hálózatokon

A Nielsen piackutató cég felmérése szerint a legtöbb időt, közel 6 óra 28 percet az olasz, míg 6 óra 25 percet az ausztrál internetezők töltötték a webes hálózatokon. Őket az amerikai felhasználók követik 6 óra 2 perces adattal. A negyedik helyen közel 5 óra 51 perces értékkel a britek, az ötödiken majd 4 óra 51 perccel a spanyolok találhatók. A hatodik helyet a brazilok (közel 4 óra 28 perc), a hetediket a franciák (4 perc 12 másodperc), a nyolcadikat (3 óra 47 perc) a németek szerezték meg. A 10-es lista utolsó két helyén a svájciak (3 óra 26 perc és meglepetésre a japánok (2 óra 37 perc) szerepelnek.

A Nielsen munkatársai arra is kíváncsiak voltak, hogy összességében mennyivel többen tartózkodnak rendszeresen a közösségi oldalakon. A tavaly februári 244,2 milliónál 30 százalékkal többen, összesen 314,5 millióan látogatták a webes hálózatokat. Az Egyesült Államokban a Facebookot és a többi hasonló oldalt látogatók száma a tavaly februári 115 millióról 149 millióra nőtt. A legtöbben a Facebookot keresik fel, a szolgáltatás az összes közösségi portállátogató 52 százalékát tudja maga mellett. A MySpace-nél ez az érték 15, a Twitternél 10, a LinkedIn-nél 6, míg a Classmates Online-nál 5 százalék.

A Facebookot átlagosan havonta 19-szer keresi fel egy felhasználó, míg a MySpace-t 6, a Twittert 5, a LinkedInt és a Classmates Online-t egyaránt 3-3 alkalommal látogatják meg az internetezők. A legtöbb időt, 5 óra 52 percet szintén a Facebookon töltik az emberek, ezt követi 59 perc 33 másodperces adatával a MySpace. A Twitternél a látogatás időtartama 36 perc 43 másodperc, a LinkedIn-nél 12 perc 47, míg a Classmates Online-nál 13 perc 55 másodperc.

/SG.hu Hírlevél, 2010. március 23., <http://www.sg.hu/>

(SzP)