



A tudományos kommunikáció elősegítése aggregált publikációs környezetekkel

Az internet forradalmi változásokat hozott a tudományos kommunikációban. 15 éve még elsősorban a helyi egyetemi könyvtár gyűjteménye határozta meg, hogy egy kutató milyen publikációkhoz férhetett hozzá. Ami ott nem volt meg, azt hosszadalmas könyvtárközi kölcsönzéssel lehetett csak beszerezni. Manapság már azonnal elérhetők a tudományos publikációk – feltéve, hogy a könyvtárnak van elég pénze előfizetni a teljes szövegű online adatbázisokra. De minthogy nem mindenhol és nem mindenre van elegendő keret, elindult az Open Access mozgalom, aminek az az alapelve, hogy a közpénzeiből finanszírozott tudományos eredmények legyenek nyilvánosan hozzáférhetők. Az elmúlt években egy sor nyílt archívum jött létre. Ezek nemcsak a kutatók közötti együttműködést könnyítik meg, hanem erősítik a tudományos világ és a társadalom közötti kapcsolatot is, hiszen így például újságírók, döntéshozók és a nagyközönség érdeklődő tagjai megismerhetik a legújabb kutatási eredményeket.

A tudományos publikálástól a tudományos kommunikációig

A kutatás során nagyon fontos a kommunikáció, és ennek csak egy része a publikálás, illetve mások publikációinak olvasása. Mielőtt egy cikk vagy könyv végleges formát ölt, a szakemberek közt rengeteg megbeszélés zajlik: nézeteket, ötleteket, véleményeket ütköztetnek – sőt esetleg a kutatási adataikat is megosztják egymással. A kiadók és az adatbázis-szolgáltatók nagyszerű eszközöket dolgoztak ki arra, hogy a szakkönyvek és szakcikkek (sőt a rájuk való hivatkozások is) könnyen és gyorsan felfedezhetőek, visszakereshetőek legyenek. A publikálási folyamatot segítő megoldások is léteznek már, de a publikációs lánc hagyományos szereplői a szélesebb értelemben vett tudományos kommunikációt nem segítik érdemben. Tőlük függetlenül alakították ki egyes szakterületek kutatói

azokat a kollaboratóriumokat (*collaboratory*), azaz együttműködési centrumokat, amelyek lehetővé teszik számukra, hogy virtuálisan szervezett kutatócsoportokban dolgozzanak együtt, közösen használjanak berendezéseket, adatbázisokat, digitális könyvtárakat vagy egyéb erőforrásokat, és természetesen közösen tudjanak így előkészíteni publikációkat is: megírni és megvitatni a kéziratokat. Jó példa minderre a komplex hálózatok vizsgálatára létrejött *Cx-Nets* (cxnets.googlepages.com) kollaboratórium, amelynek keretében amerikai, olasz és francia kutatócsoportok dolgoznak együtt. Az így megosztott információ azonban nem feltétlenül nyilvános; vagyis bár a kollaboratóriumok a modern tudományos kommunikáció minden jellemzőjével rendelkeznek, nem a nyílt hozzáférés elősegítése az elsődleges céljuk.

Értéknövelt publikációk és aggregált publikációs környezetek

Létezik egy olyan új publikációs és kommunikációs modell, amely egyesíti a hagyományos tudományos közlemények jellemzőit a kollaboratóriumok kibővített lehetőségeivel. Megteremti annak a feltételeit, hogy egy publikációt további információs objektumok (adatok, képek, hangok, vélemények stb.) egészítsenek ki, egyúttal leírja ezek kapcsolatát is a „tudományos publikációs csomagon” belül. Ez a fajta „értéknövelt” publikáció (*EP = Enhanced Publication*) nem egy statikus és lezárt dokumentum, mint amilyen egy folyóiratcikk, hanem dinamikus: idővel változhat, bővíthet, megjegyzések fűzhetőek hozzá, sőt egyes elemei felhasználhatók más EP-k vagy hagyományos publikációk részeként. Az ilyen komplex publikációk létrehozása technikailag nem egyszerű. Szükség van olyan eszközökre, amelyek megkönnyítik a kutatóknak az információs objektumok összekapcsolását és a köztük levő relációk leírását. Ezt a feladatot segíti az ún. aggregált publikációs környezet (*APE =*

Aggregated Publications Environment), amely az OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*) modellre épül.

Az OAI-ORE specifikáció webes források csoportjainak leírását és egyben ezek lehatárolását teszi lehetővé szabványos, géppel értelmezhető formában. Ilyen csoport lehetnek például: egy képgyűjtemény, egy könyv fejezetei, egy tanulmány a hozzá kapcsolódó adatsorokkal. (Részletesebben I. erről *Rumsey* és *O'Steen* cikkének bevezetőjét [1].) A modell három fő elemet különböztet meg: a forrásokat (*Aggregated Resources*), a köztük levő kapcsolatokat leíró forrástérképet (*Resource Map*) és magát az aggregációt (*Aggregation*), vagyis az így lehatárolt csoportot (1. ábra).

Az OAI-ORE modell könnyen felhasználható EP-k leírására. Az aggregáció az EP-t alkotó információs objektumokból áll össze, és ezek egymáshoz való viszonyát a forrástérkép reprezentálja. Nemcsak EP-k írhatók így le, hanem akár hagyományos publikációk is, mivel valójában ezek is több részből állnak: például *Bevezetés*, *Módszerek*, *Eredmények*, *Irodalom* – egy így felosztott, majd újra egybefoglalt tanulmánynak az az előnye, hogy egyes részei önállóan is hivatkozhatók vagy beépíthetők más dokumentumokba. Hasonló moduláris publikációkat javasolt egyébként az *Elsevier* kiadó is [3].

Az ESCAPE projekt

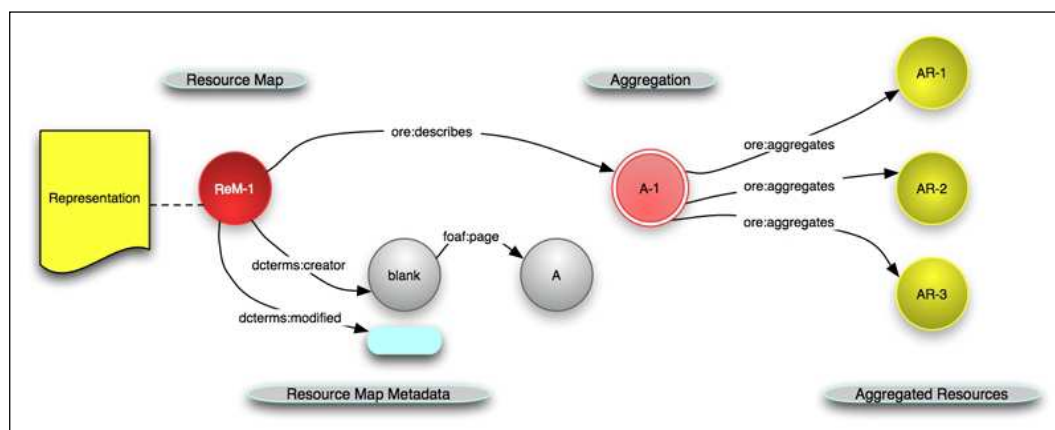
2008-ban a holland felsőoktatási és kutatási alap, a *SURF* foundation tendert írt ki EP környezetek kialakítására. Ennek egyik nyertese az *ESCAPE* (*Enhanced Scientific Communication by Aggregated Publications Environments*) nevű projekt lett (escapesurf.wordpress.com). Az *Universiteit Twente* és a *Rijksuniversiteit Groningen* tudósai és kutatói dolgoznak együtt informatikusokkal, programozókkal és könyvtárosokkal, hogy létrehozzanak egy olyan számítógépes környezetet, amelyben könnyen lehet ilyen értéknövelt publikációkat alkotni és megosztani. A világ más részein is léteznek már ehhez hasonló kezdeményezések, mint például a *Max Planck Society eSCIDoC* projektje (különbéle információs objektumok összekapcsolása egyetlen tudományos tudástérbe), vagy a *SCOPE* nevű ausztrál fejlesztés (szerkesztőeszközök EP-k számára), illetve az *ICE-TheOREM* projekt (OAI-ORE funkciók beépítése egy tartalomkezelő rendszerbe).

A 2009 decemberéig tartó ESCAPE projekt keretében három kutatócsoport számára készült aggregált publikációs környezet: 1. a konfliktus, a kockázat és a biztonság témájával foglalkozó *iCRiSP* központ, 2. az ultra-nagysebességű kamera alkalmazásait fejlesztő *Brandaris 128* projekt, és 3. a közrend és a közbiztonság témájában kutató *Centrum voor Openbare Orde en Veiligheid* részére folyik ilyen APE-k fejlesztése. Az ezekben dolgozó kutatóknak olyan eszközökre van szükségük, amelyek nem igénylik az OAI-ORE modell és az azt megvalósító RDF nyelv ismeretét. Fontos az is, hogy ne kelljen lényegesen módosítani a már meglévő repozitóriumokhoz ahhoz, hogy az aggregált dokumentumokat és azok metaadatait is tárolni és keresni lehessen velük. Mivel az EP-knek – pontosabban az azok szerkezetét leíró forrástérképeknek – is saját egyedi azonosítójuk (*URI = Uniform Resource Identifier*) van, ezért ezeket ugyanúgy egyetlen információs egységként lehet kezelni, mint a szokásos (pl. egyetlen PDF állományból álló) publikációkat.

Meg kell oldani továbbá a hozzáférési jogosultságok kérdését is, hiszen az EP a kooperáció új formáinak nyit teret, de természetesen az nem lenne szerencsés, ha egy kutatócsoporton belül mindenki szabadon módosíthatná az EP-t alkotó objektumokat. Ezért az ESCAPE projektben három hozzáférési szintet definiáltak:

1. A szerzőknek teljes jogkörük van, szabadon bővíthetik, törölhetik vagy módosíthatják a publikációt alkotó információs objektumokat; továbbá kijelölhetnek „szerkesztőket”.
2. Ezek a szerkesztők hozzáadni tudnak (pl. címkézhetnek vagy kommentezhetnek egyes részeket), de nincs joguk módosítani vagy törölni.
3. Mindenki más csak megjeleníteni tudja az EP-t (illetve annak a forrástérképét), de nincs módosítási lehetősége.

Első lépésként létrehozták a forrástérképek archívumát, erre a célra a *Fedora 3.1* verzióját választották, leginkább azért, mert ez a repozitóriumkezelő rendszer támogatja az RDF tripleteket (az „alany-állítmány-tárgy” típusú állításokat, amelyeket az objektumok közötti kapcsolatok leírására használnak). Ezeknek a forrástérképeknek a metaadatai ugyanúgy begyűjthetők a szokásos OAI-PMH protokollon keresztül, mint a hagyományos, egyedi digitális dokumentumok adatai. Így, ha egy felhasználó lekér egy értéknövelt publikációt a repozitóriumból, akkor egymáshoz kapcsolt objektumok hálózata jelenik meg a képernyőjén. Sőt, az OAI-ORE modellben definiált ún. *proxy URI*



1. ábra Az A-1 aggregációt, amely három forrást (AR-1, AR-2, AR-3) kapcsol össze, a ReM-1 forrástérkép írja le. [2]

segítségével arra is van lehetőség, hogy egy adott információs objektumot adott kontextussal (aggregációs környezetben) együtt kérjünk le az archívumból (pl. egy bizonyos kiállítás egyik képét annak leírásával együtt).

A forrástérképek megalkotásához és módosításához az ESCAPE projekt fejlesztői egy felhasználóbarát térképszerkesztőt készítettek. Ennek segítségével az arra jogosult szakemberek relációkat hozhatnak létre és írhatnak le az objektumok között. A jövőben arra is mód lesz, hogy a felhasználók is hozzáadjanak szabadon választott leírásokat ezekhez a kapcsolatokhoz. Az EP-eket alkotó objektumok közötti relációk típusait – amennyire lehet – a már létező, elterjedt szótárakból igyekeztek összeválogatni, ezzel is megkönnyítve a későbbi együttműködést más rendszerekkel. De mivel sokféle kapcsolat lehet információs objektumok között, és ráadásul ezeket a különböző tudományterületek szakemberei különbözőképpen írják le, az ESCAPE keretében fejlesztett publikációs környezetben lehetőséget teremtettek arra, hogy a szokásos *dcterms:relation* metaadatot egy *relation-Annotation* nevű leírással is kiegészíthessék, ahol is szabadon meg lehet adni, hogy pontosan milyen kapcsolatban van az egyik objektum a másikkal.

Tanulságok

A szemantikus web és az OAI-ORE koncepcióban komoly tudományos lehetőségek vannak. Amellett, hogy minden objektumra saját URI-val lehet hivatkozni, ezek kapcsolatai is leírhatók és ábrázolhatók egy – szintén saját azonosítóval rendelkező –

forrástérképen, s ennek határai is beállíthatók. A szerzők így összekapcsolhatják azokat az elemeket, amelyeket a publikációjuk szempontjából fontosnak tartanak, az olvasók számára pedig világos lesz, hogy mik képezték az illető közlemény alapját. Vagyis az EP típusú publikálás „átláthatóbbá” teszi a tudományos kutatást, várhatóan gyorsítja is azt és közelebb hozza a tudományt a társadalomhoz.

Ezek az új fejlemények azonban új problémákat is felvetnek. Először is a forrástérképek megalkotása jóval bonyolultabb feladat, mint egyszerűen feltölteni egy publikációt egy intézményi repozitóriumba. A kutatóknak jól kell ismernie az aggregálni kívánt információs anyagokat és a köztük létrehozható kapcsolatokat. Ehhez „felülről” kell rálátnia az egész információhalmazra.

Továbbá vannak bizonyos gyakorlati – remélhetően csak átmeneti – nehézségek is: a felület nem elég felhasználóbarát, a relációk leírására használt szótárak nem teljesek, és nem lehet még visszakeresni, hogy egy adott objektum milyen forrástérképeken szerepel. Utóbbihoz be kellene vezetni egy olyan *ping* vagy *trackback* funkciót, amely minden alkalommal, amikor valaki felvesz egy objektumot egy EP forrástérképre, visszaküldi annak (és persze az adott objektumnak) az URI-ját annak a repozitóriumnak, amely az illető objektumot őrzi.

Vagy ott van például a névazonosítók kérdése: egy EP típusú publikációs környezetben minden objektumnál meg kell adni egy URI-t, amely az objektum készítőjét azonosítja. De sok esetben egy embernek több ilyen azonosítója is van a weben (pl. van

egy oldala az egyetemi szerveren és egy másik annak a kerékpáros klubnak a honlapján, amelynek tagja); hogy mikor melyiket célszerű választani, az helyzetfüggő. Vagyis ilyenkor nagy szükség lenne az egységes digitális szerzői azonosítókra. Mindezen problémáktól függetlenül már most is vannak konkrét gyakorlati alkalmazások: A *Dataplus* projektben például a kutatók hagyományos típusú publikációkat kötnek össze kutatási adatsorokkal. Az iCRISP központban a tudósok és más illetékes szakemberek a társadalom tagjainak a vízminőséggel kapcsolatos kockázatészlelésére és adaptív viselkedésére vonatkozó ismereteiket osztják meg egymással, az ESCAPE forrástérkép szerkesztőjét használva. Szeretnék ösztönözni ezzel is a tudományos kutatók és a politikacsinálók közötti információcserét. Egy harmadik esetben pedig az ügyvédek tudnak létrehozni olyan EP-eket, amelyekben a hatályos joganyaghoz magyarázatok, fogalomdefiníciók, jegyzőkönyvek és konkrét jogesetek is kapcsolódhatnak.

Irodalom

- [1] RUMSEY, S. – O’STEEN, B.: OAI-ORE, PRESERV2 and Digital Preservation. = *Ariadne*, 57. sz. 2008. október 30.
<http://www.ariadne.ac.uk/issue57/rumsey-osteen/>
 - [2] LAGOZE, C. – Van de SOMPEL, H.: Ore User Guide. = Primer, 2008. október 17.
<http://www.openarchives.org/ore/1.0/primer>
 - [3] Cell Press beta, „Article of the Future”, 2009. július 20.
<http://beta.cell.com/index.php/2009/07/article-of-the-future/>
- /HOGENAAR, Arjan: Enhancing Scientific Communication through Aggregated Publications Environments. = *Ariadne*, 61. sz. 2009. október 30.**
<http://www.ariadne.ac.uk/issue61/hogenaar/>

(Drótos László)