

Az internet tűnékenysége és az oktatás

(A hazai internet-archiválás hiányának hatása az oktatásra)¹

Dr. Kokas Károly informatikai és információs főigazgató-helyettes
SZTE Klebelsberg Könyvtár
<kokas@ek.szte.hu>

A nyilvános hozzáférésű weboldalakon megjelenő információ és tudás az oktatásban, a kutatásban és a nemzeti emlékezet szempontjából is hasonló fontosságúvá vált, mint ezek hagyományos, papíron közzétett formái. A könyvben, folyóiratban és egyéb kiadványokban levő ismeret a könyvtárakban hosszú ideig (akár évszázadokig) is megmarad: kikereshető, megismerhető, idézhető... A világhálós oldalak viszont bármelyik pillanatban eltűnhetnek vagy megváltozhat a tartalmuk. Az online források változékonysága megnehezíti a beépülésüket a tudományos és az oktatási tevékenységbe, hiszen mindkét területen fontos, hogy a források évtizedekig biztosan elérhetőek maradjanak és lehessen rájuk hivatkozni.

1. Az internet-archiválás szükségessége²

Talán azt ma már, hogy miért fontos a digitális kultúra megőrzése, nem kell nagyon megmagyarázni. Mindnyájan napi élményként éljük meg, hogy mennyire meghatározó lett az utóbbi pár évtizedben a világhálózat, az internet. Tudjuk, hogy mennyi minden már csak ott érhető el; és látjuk, érezzük azt is, hogy milyen óriási tempóban növekszik a világ és az egyes nemzeti kultúrák világhálós „tartalomipari termése” – és azt is látjuk, hogy mindez rohamosan pusztul is. Nem pusztán arról van szó, hogy egy-egy rávezető (link) elvész, egy idő után nem vezet sehová, hanem egész világhálós oldalak, sőt honlapok ezrei tűnnek el nyomtalanul egyik napról a másikra, mintha a nagy „digitális Moloch” felzabálná gyermekeit, szinte ugyanolyan tempóban, ahogy megszülettek.

Másrészt az is igaz, hogy a digitális világ jelentős részének napi „megsemmisülése” – úgy látszik – ma még nem vagy kevésbé lépi át az ingerküszöbünket. Emlékezhetünk rá, ha a rossz tárolási körülmények miatt meg kellett semmisíteni a Nemzeti Könyvtár egyik tartalékolt kisnyomtatvány-gyűjteményi részét, vagy valahol csőtörés miatt eláztak régi folyóiratok, az milyen kemény kritikai visszhangot váltott ki a médiában. A digitális anyagok eltűnése, noha ez mindennapos eset, jóval kisebb hatást kelt, megmentésüket és megőrzésüket senki nem érzi feladatának. Pontosabban szólva: nálunk nem, mert a világ más részein már elég sok munkát elvégeztek ezen a téren, és – bár a feladat kétségtelenül nagyon nehéz – vannak már biztató eredményeik is, sőt illetően szolgáltatások is elindultak.

A kívánság hazai megfogalmazása pedig elég régi, s van hozzá tartozó „politikai” kezdeményezés is, hiszen az UNESCO Közgyűlésének 32. ülésén, 2003. október 17-én elfogadta a „Charta a digitális örökség védelméről”³ című határozatot, az pedig világosan

¹ A dolgozat megírásához felhasználtam Drótos László és Moldován István (OSZK) különféle anyagait, ill. a velük közösen készített szakmai összeállításokat.

² Vö. Drótos László gondolatmenetével: <http://mek.oszk.hu/html/irattar/eloadas/2006/mia.htm>

³ <http://www.unesco.hu/index.php?type=node&id=508>

megfogalmazza az ezzel kapcsolatos feladatokat, és külön kiemeli, hogy a válogatott archiválás esetén a már eleve digitálisan született anyagoknak elsőbbséget kell adni. Ez a dokumentum meghatározza a felelősöket is: „*A digitális örökség megőrzése a kormányok, alkotók, kiadók, releváns iparágak és az örökségvédelmi intézmények kitartó erőfeszítéseit igényli.*”⁴

2. A webarchiválás és az oktatás⁵

A felsőoktatás értelemszerűen számtalan változatban és módon kutatja, az oktatás teljes egésze közvetlenül felhasználja a kulturális javak sokaságát. Gondoljuk végig, hol és hányféleképpen bukkan föl a mindig és jól elérhető világhálós forrás szükségessége az oktatás különféle területein?

- A nyilvános hozzáférésű weboldalakon megjelenő információ és tudás az oktatásban hasonló fontosságúvá vált, mint ezek hagyományos, papíron közölt formái. A könyvben, folyóiratban és egyéb kiadványokban levő ismeret a könyvtárakban hosszú ideig (akár évszázadokig) megmarad: kikereshető, megismerhető, idézhető, hivatkozni lehet rá. Mint említettük, a világhálós oldalak viszont bármelyik pillanatban eltűnhetnek vagy teljesen megváltozhat a tartalmuk. Az online források változékonysága megnehezíti a tudományos és oktatási tevékenységbe való beépülésüket, hiszen mindkét területen fontos, hogy az információk évtizedekig biztosan elérhetők maradjanak és utalni lehessen rájuk.
- A magyar internet-archiválási tervben meghatározott gyűjtőkör a magyar vagy magyar vonatkozású tudományos, oktatási, kulturális és közéleti oldalakra terjedne majd ki, ezekből kívánna minél többet rendszeresen eltárolni és nyilvánosan hozzáférhetővé tenni. Ha az archivált anyagot, illetve annak egyes részeit stabil azonosítókkal látnánk el, akkor az egyes weboldalakra, illetve azok elemeire (pl. ábrákra, táblázatokra) hosszú távon is megbízhatóan lehetne hivatkozni. Vagyis az archivált anyag mind az oktatóknak és tudományos kutatóknak, mind pedig a diákoknak megbízhatóan használható információ- és tudásbázist jelentene. Ez óriási előny volna az ilyen típusú anyagok világhálós oktatási rendszerekben, (táv)oktatási csomagokban tervezett alkalmazását illetően, hiszen hosszú időre megadná az ismeretek eredeti összefüggésükben való kikereshetőségének lehetőségét és a rájuk való hivatkozás megbízhatóságát.
- A felsőbb fokú oktatás értelemszerűen számtalan változatban s módon kutatja s közvetlenül fel is használja a kulturális javak sokaságát. Ennek egy igen lényeges vetülete a múlt bármely korszakára vonatkozóan az adott kultúra, korstílus, divat megjelenése, látványterve, stílusbeli jellemzői. Ez adott esetben rendkívül gyorsan változhat, az elektronikus megjelenésű forrásoknál akár naponta is. Mindnyájunk számára értékes forrása a múlt megjelenési formáinak egy folyóirat címlapja, egy

⁴ Drótos László: Mi a MIA? – Javaslat egy Magyar Internet Archivum létrehozására = Tudományos és Műszaki Tájékoztatás, 2006. 6. szám. http://tmt.omikk.bme.hu/show_news.html?id=4431&issue_id=473

⁵ Kokas Károly: Könyvtárak a Rubiconnál = Educatio, 2013. 3. szám 363-376. ill. http://www.hier.iif.hu/hu/educatio_reszletes.php?id=96

könyvborító, egy plakát stb., sőt ezeknek művészeit is jól ismerjük, s nem egy esetben a kor egész stílustörténeti besorolását illetően meghatározónak tartjuk ezeket. A webkultúra esetében azonban archívum nélkül a múlt ezen rétegei a tartalommal együtt nyom nélkül eltűnhetnek.

- Egy magyar webarchívum azonban nemcsak azt tenné lehetővé, hogy megnézhető legyen egy időközben már megszűnt vagy megváltozott világhálós oldal, hanem az archívumra ráépülő, értéknövelt szolgáltatások által az „élő” weben lehetségeseknél fejlettebb keresések és vizsgálatok is elvégezhetőek lennének rajta. Így pl. adat- és szövegbányászat, az idődimenziót is figyelembe vevő statisztikai és nyelvi elemzések sokasága stb. Ennek a magyar nyelv- és stílustörténet kutatásától a kommunikáció oktatásán keresztül a szociolingvisztikáig igen széles a tanításban, kutatásban való hasznosíthatósága. Nyilvánvaló, hogy mindez kapcsolatban van a nyelv- és beszédtechnológiai trendekkel és következményeikkel (pl. mint óriási „adatsilója” a mindennapi nyelvhasználat különféle területeinek), valamint a társadalomtudományok igényeivel (mind szakadatbázis, mind archiválás szintjén), és a humántudományok infrastrukturális kezdeményezéseivel.
- Nem mellékes nyeresége egy archívumnak, hogy automatikusan megoldódik vele az intézmények honlapjainak archiválása. Jelenleg a legtöbb intézmény (pl. iskola, egyetem, főiskola és kutatóintézet) nem archiválja honlapjának tartalmát, vagy pedig csak rövid ideig. Könnyebb és helytakarékosabb az intézmény honlapját folyamatosan alakítgatni, és legfeljebb egy komolyabb arculatváltáskor őrzik meg egy ideig az előző verziót. Így viszont az intézmény múltjának dokumentálásából mind jelentősebb hányad vész el, hiszen egyre több minden már csak az internetre kerül fel, papírra nem is nyomtatják ki. Egy webarchívum képes lehetne az egész oktatási intézményrendszer számára ún. „digitális levéltári” funkciót is nyújtani, különösen akkor, ha az intézmények eleve, vagy pedig az archiválás után ellátják megfelelő részletességű metaadatokkal a fontosabb dokumentumaikat a későbbi visszakeresés megkönnyítése érdekében.
- Részletkérdésnek tűnhet, és talán az oktatás szempontjából sokadlagos, de adott esetben rendkívüli fontosságú lehet egy egészen más aspektusa is a magyar webarchiválásnak: az, hogy jogérvényesítő vagy ilyen-olyan értelmezési viták esetén a hivatalos honlapok időtengely mentén precízen metaadatolt változatai *független harmadik féltől* szerezhetőek be, amely garantálni tudja az adott változat időbeli valóságos s hiteles voltát. Így az archívum hiteles helyként szolgálhatná az oktatást, a kutatást és az egész múltmegőrzést is.

Összefoglalva elmondhatjuk, hogy a magyar internet-archiválás lehetővé tenné az oktatási anyagok fejlesztését és integrálását, és nagy mértékben segítené az internetes adatvagyon archiválását, valamint az oktatásban és kutatásban való hasznosítását, összhangban az egész életen át tartó tanulás gondolatával is, mely függetlenül attól, hogy divat hangoztatni, nagyon is igaz

3. Külföldi példák

A világban több tucat ilyen, nagy nyilvánosságnak elérhető szolgáltatás létezik, ezek jellegükben és méretükben is erősen különbözőek. Jellemző, hogy sok köztük ma még inkább

csak pilot-jellegű próbálkozás, amelyek csupán egy előre lehatárolt, néhány éves időszakra terjednek ki, kevés a már üzemszerűen működő, valóban nagy archívum. Ám jellemző az archívumok integrálódása is, a közös szabványokra, technológiákra és a munkamegosztásra való törekvés⁶.

A legrégebb és legismertebb kezdeményezés az *Internet Archive*⁷ nevű nonprofit szervezet, San Franciscóban. Ez a webarchiváláson túlmenve valósággal a digitális objektumok Alexandriai Könyvtárává szeretne válni. Legismertebb szolgáltatásuk talán az ún. Wayback Machine, ez egy 1996-tól fogva keletkezett weblapokból összegyűjtött anyagot tartalmaz, mintegy 250 milliárdnyi (!) weboldalt tud előkeresni⁸, rámutató (URL) alapján. Újabb fejlemény a teljes szövegű kereső beépítése. A „webaratás” elvileg a web nyilvános részének egészére kiterjed, a népszerűbb oldalakat gyakrabban gyűjtik be (az átlagos gyakoriság 1-2 hónap, az állomány átlagos növekedése ma már több tucat terabyte havonta), az anyag később nyilvánossá válik, de az archívum némely részei csak kutatóknak érhetők el. Mindazonáltal még ezt a rendszert sem tekinthetjük valamiféle „világ-webarchívumnak”, mert számtalan korlátja van, s óriási méretei ellenére korántsem teljes. Így a tematikus és nemzeti webarchívumok előtt még mindig komoly perspektíva áll.

Az említett wiki szócikk listája⁹ egyébként legalább 60 nagyobb projektet analizál a világ minden részéről, az egyes jelentősebb vállalkozásokhoz külön link visz el. Pl. a cseh webarchívum idén 14 éves, 526 millió fájl tartalmaz és mindennek tárolása 24 TB helyet foglal el¹⁰. De a kis nemzetek közül említhetnénk még Dániát is, ahol a *Netarkivet.dk* helyen¹¹ a nemzeti webarchiválási program keretében 16 milliárd fájl (!) archiváltak eddig és fél petabájt (0.5 PT) helyen tárolják mindezt.

4. Hazai terv az archiválásra

2006-tól – a MEK-alapítója, Drótos László javaslatára (Drótos, 2006) – folyamatosan készült a Magyar Internet Archívum (MIA) tervezete, és ma már eléggé ki van dolgozva a megvalósításhoz. A Magyar Internet Archívum tervezett gyűjtőköre a magyar vagy magyar vonatkozású, tudományos, oktatási, kulturális és közéleti oldalakra terjed majd ki, ezekből kíván minél többet rendszeresen lementeni és nyilvánosan hozzáférhetővé tenni. Az archivált anyagot, illetve annak egyes részeit stabil azonosítókkal látják el, így az elmentett weblapokra, illetve azok elemeire hosszú távon is megbízhatóan lehet majd hivatkozni. Vagyis az archivált anyag a felhasználó magánszemélyek és intézmények számára is megbízhatóan használható információ- és tudásbázist jelent majd.

⁶ Bővebben ezekről: http://en.wikipedia.org/wiki/Web_archiving

⁷ <http://www.archive.org>

⁸ http://en.wikipedia.org/wiki/Wayback_Machine

⁹ Wikipédia szócikk: http://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives A táblázatok nem csak felsorolják, de metódus, méret, visszakereshetőség szerint is adatolják a jelentősebb projekteket.

¹⁰ <http://www.webarchiv.cz/>

¹¹ <http://netarkivet.dk/in-english/>

A hálózati infrastruktúra, a tárhely adva lehet (NIIF), a szakmai irányítás szerve szintén (OSZK), s van néhány nagy egyetemi műhely, amely tudna szakmai támogatást adni (Szeged, Debrecen stb.). Bizonyos értelemben van kidolgozott kezdő „menetrend”, sok minta áll rendelkezésre. Kétségtelen azonban, hogy a tervezetet sajnos még mindig értetlenkedés fogadja, fantazmagóriának tekintik, vagy mert a szakmán kívüli döntéshozók nem ismerik a nemzetközi trendeket, vagy mert sokan nem érzékelik az erodálódás mértékét. Sajnos a közelmúltban több nagy terv megbukott, nincs támogatás. Itt bizony elmaradásban vagyunk, de a közelmúltban talán valami megindult, hogy e kétségtelen hiányterületen egy működő projekt és mögé megbízható intézményi háttér, sőt szabályozó törvényi előírás is kerüljön.¹²

¹² A kormányzat tervezi, hogy miniszteri rendelet útján vagy más módon inicializálja a témát.