noWebarchive

Holl András - MTA KIK

404 Not Found – Ki őrzi meg az internetet?
OSZK
2019 Nov. 14.

Gyakorlatból tudjuk, hogy a web egyre fontosabb az információkat közvetítő médiumok között, és napról-napra információt vesztünk a weboldalak megszűnésével és átalakulásával. A web dinamikus, a web centralizált (egy adott információ létezhet csak egyetlen szerveren - ellentétben a nyomtatott dokumentumok által biztosított redundanciával).

Webarchiválásra szükség van - ezt itt nem kell bizonyítanom. Ám egy logikus, átgondolt világban az információkat másképpen kellene megőrizni. Figyelemmel kellene lenni arra, hogy a friss információk ne töröljék a régi (ám érvényes) információkat. Ez sok webhelyen (részben) így is működik: a hírek mellett például létezik egy hír archívum. A másik probléma az, hogy a web technikailag nehezen archiválható, például a média lényegéhez tartozó hiperhivatkozások miatt, részben azért, mert a technológia gyorsan változik, a régi weboldalakat egy újabb böngésző nem feltétlenül jeleníti meg helyesen. És természetesen megoldást kell találni a problémára, hogy a gazdátlanná vagy elavulttá vállt weboldalak ne tűnjenek el nyom nélkül.

Amennyiben az információ szolgáltató számára az információk hosszú távú megőrzése fontos, és tisztában van a webarchiválás nehézségeivel, más stratégiát is alkalmazhat az információk megőrzésére. Elfogadhatja azt az alapelvet, hogy a web csak kirakat, az adatoknak másutt, más, könnyebben megőrizhető formában is meg kell lenniük. Két utat is bemutatunk az alternatív információ megőrzésre.

A.) A "kirakat" és az "adattár" szétválasztása

Példák az MTA KIK valós vagy megvalósítani kívánt gyakorlatából.

"Hazádnak rendületlenül - 180 éves a Szózat"

Kiállítás Vörösmarty Mihály hagyatékából / Irodalmi adatbázis

Keresés cím szerint

Szózat

Keresés álnév szerint

Az álnév néhány karaktere

Keresés tárgyszó szerint

A tárgyszó néhány karaktere

Keresés folyóirat szerint

A folyóirat címének néhány karaktere

Keresés megjelenés éve szerint

A megjelenés évszáma

Keresés

KERESÉSI ÚTMUTATÓ

A Magyar Tudományos Akadémia Könyvtár és Információs Központ Szakinformatikai Osztályának gondozásában most megjelenő irodalmi adatbázis, a virtuális kiállítás tematikájához kapcsolódóan, Vörösmarty életművéből közel 300 művet ad közre: a válogatás költői műveket, elbeszéléseket, drámákat és vegyes prózai dolgozatokat tartalmaz. A korabeli irodalmi folyóiratok közül csak azokat dolgoztuk fel, amelyek digitalizált változatai megtalálhatók könyvtárunk repozitóriumi gyűjteményeiben. E teljes szövegű adatbázis tehát irodalmi forrásértékű, nyomtatott, első szövegközlések gyűjteménye, sőt, amikor nem maradt fent a szerzői kézirat, Vörösmarty műveinek elsődleges, szöveghű forrásait mutatjuk be, digitális formában.

A Magyar Tudományos Akadémia Könyvtár és Információs Központ Repozitóriuma (Repository of the Academy's Library - REAL) a könyvtár olyan információs tere, amely a tudományos információt teljes szöveggel, digitális formában, nyílt forráskódú, szabad szoftver (Eprints) támogatásával archiválja. Az így tárolt tudományos információt metaadatokkal feltárja, visszakereshetővé teszi és a nyílt hozzáférés (Open Access) elveit követve kínálja az olvasóknak.

Tovább >>

Sallai-Tóth László
szerkesztő

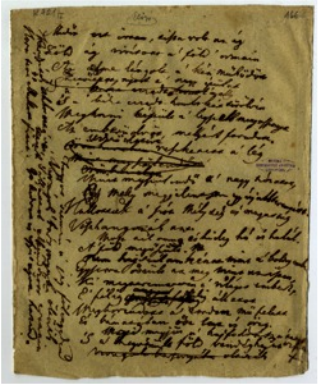
Date Deposited: 15 Jun 2017 13:45

Last Modified: 15 Jun 2017 13:47

URI: http://real-ms.mtak.hu/id/eprint/15772

Actions (login required)

View Item



Preview

Előszó

Vörösmarty, Mihály (1850) *Előszó.* , H.n. (Manuscript)

PDF
Vorosmarty_Eloszo_K721_1_166.pdf
Download (21MB) | Preview

PDF
Vorosmarty_Eloszo_K721_1_167.pdf
Download (16MB) | Preview

PDF
Vorosmarty_Eloszo_K721_1_167a.pdf
Download (15MB) | Preview

Item Type: Other

Additional Information: Jelzet: K 721/I.166-167.a. 166. „Midőn ezt írtam...” [Előszó, fogalmazvány] 167-167.a. Előszó. [Tisztázat] Autogr. 3 f.

Subjects: P Language and Literature > PN Literature (General)

SWORD Depositor: ALEPH SWORD

Depositing User: Béla Mázi

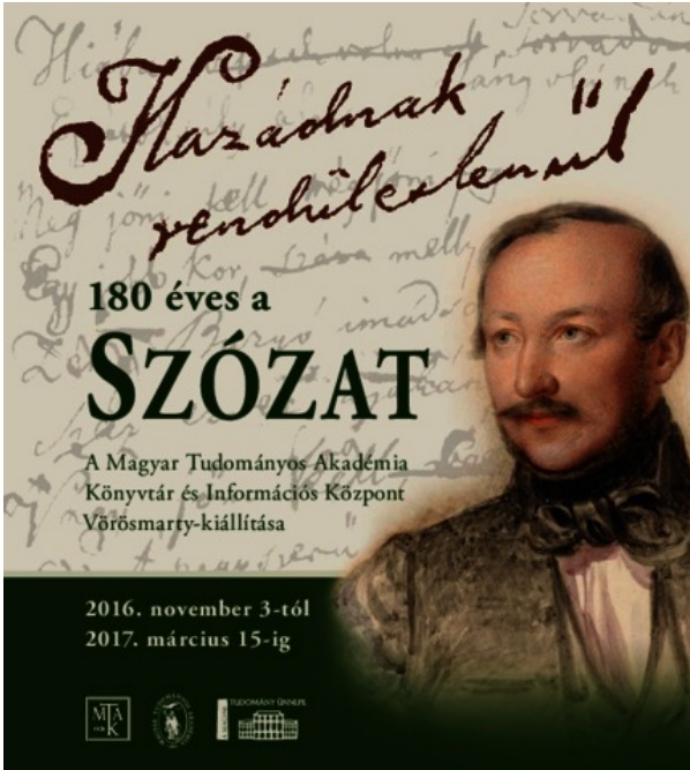
Date Deposited: 15 Jun 2017 13:45

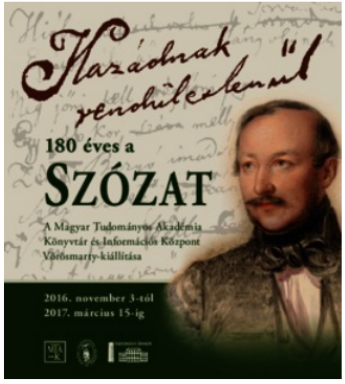
Last Modified: 15 Jun 2017 13:47

URI: http://real-ms.mtak.hu/id/eprint/15772

Actions (login required)

View Item


Hazádnak vendülésemül
180 éves a
SZÓZAT
A Magyar Tudományos Akadémia
Könyvtár és Információs Központ
Vörösmarty-kiállítása
2016. november 3-tól
2017. március 15-ig



Preview

180 éves a Szózat : a Magyar Tudományos Akadémia Könyvtár és Információs Központ Vörösmarty-kiállítása

Babus, Antal (2016) *180 éves a Szózat : a Magyar Tudományos Akadémia Könyvtár és Információs Központ Vörösmarty-kiállítása.* Magyar Tudományos Akadémia Könyvtár és Információs Központ, Budapest.

Text
Kisero fuzet 1017.pdf
Download (5MB) | Preview

Item Type:	Book
Subjects:	D History General and Old World / történelem > D4 Modern History / újkor története D History General and Old World / történelem > DN Middle Europe / Közép-Európa > DN1 Hungary / Magyarország P Language and Literature / nyelvészet és irodalom > PH Finno-Ugrian, Basque languages and literatures / finnugor és baszk nyelvek és irodalom > PH04 Hungarian language and literature / magyar nyelv és irodalom
SWORD Depositor:	ALEPH SWORD
Depositing User:	Erika Bilicsi
Date Deposited:	26 May 2017 12:38
Last Modified:	13 Jun 2017 11:38
URI:	http://real-eod.mtak.hu/id/eprint/4245

Actions (login required)

A személyesen meglátogatható, fizikai térben történő kiállítás meghatározott ideig volt nyitva. Volt azonban két, tartósabb inkarnációja: a weblap és a kísérő füzet. A weblap részben aktuális információkat (pl. nyitvatartás) közölt, illetve a zárás után is fenntartotta a virtuális látogathatóság lehetőségét. A webes felület ma is létezik, lehet, hogy évek múlva is létezni fog - de vajon szükséges-e archiválni? Nem biztos, mivel a kísérő füzet a weboldal megszűntével is rendelkezésre áll majd. Nyomtatott változatban, és digitálisan is: http://real-eod.mtak.hu/4245/ . Ha jól végeztük dolgunkat, a kiállítás megfelelően dokumentálva lesz: a tudományos hátteret alaposan kifejtő publikációkban, a kísérő füzetben, a könyvtár éves jelentésében. A kiállításon bemutatott dokumentumok digitális másolatai pedig elérhetőek maradnak a repozitóriumban.

9.2 Kiállítások

Hazádnak rendületlenül... Kiállítás Vörösmarty Mihálynak az MTA Könyvtár és Információs Központban őrzött hagyatékából

A Magyar Tudományos Akadémia elnöke és az MTA Könyvtár és Információs Központ főigazgatója védnöksége alatt a Magyar Tudomány Ünnepe 2016. évi rendezvénysorozata keretében rendeztük meg Vörösmarty Mihálynak az MTA Könyvtár és Információs Központban őrzött hagyatékából rendezett kiállítását. A megnyitón bemutattuk a **Szózat hasonmás kiadását**, amely a Kossuth Kiadóval való együttműködésben jelent meg. A megnyitón köszöntőt mondott **Török Ádám**, az MTA főtitkára, a Magyar Tudomány Ünnepe 2016. évi rendezvénysorozatát **Barnabás Beáta** nyitotta meg, az MTA főtitkárhelyettese, a Magyar Tudomány Ünnepe Programtanácsának elnöke, a kiállítást megnyitotta és a Szózat hasonmás kiadását bemutatta **Monok István**, a MTA Könyvtár és Információs Központ főigazgatója, a szózatról előadást tartott **Szilágyi Márton**, az ELTE egyetemi tanára, a kiállítást bemutatta **Babus Antal**, az MTA KIK Kézirattár és Régi Könyvek Gyűjteménye vezetője.

A megnyitó időpont: 2016. november 3, 12:30 óra, résztvevők száma: 150 fő.

A kiállítás 2016. november 3-tól 2017. március 15-ig tekinthető meg. Helyszín: MTA Könyvtár és Információs Központ, Budapest, V. kerület, Arany János u. 1. Nyitva tartás: hétfő–péntek 10.00–18.00, szombat 10.00–16.00. A belépés díjtalan. 2016-ban 734 fő tekintette meg a kiállítást. A kiállítást Babus Antal és Horányi Károly rendezte, és Imre Olga (Stáb-Studio Kft.) kivitelezte.

Lovász László MTA elnök a Szózat kiállítás megnyitóján

A "könyv" - még ha digitális könyv is - kiválóan alkalmas lehet a hosszú távú megőrzésre. A digitális kópia - például PDF/A formátumban - megőrizhető, kereshetőséget biztosít, akár ki is nyomtatható. Az intézményeknek a tevékenységüket, az általuk nyilvánosságra hozott információkat lehetőség szerint könyv formában is elő kell állítaniuk: mint a kiállítás kísérőfüzete, mint egy kiállítás-katalógus, mint egy évkönyv.

Az egész elképzelést talán a legjobb a feje tetejére állítani. Nem a web tartalmát célszerű áttenni könyv formába, hanem bizonyos tartalmakat adatbázisban kell tárolni, és mind a webes felületen megjelenő információt, mind az évkönyvet, intézményi telefonkönyvet, az adatbázisból dinamikusan képezni! Nem pusztán CMS szoftver alkalmazásáról beszélek, hanem a háttér-adatbázisok megfelelő szerepéről. Az általam felvázolt weboldal lehet nem CMS alapú is, amelyik pusztán meghív és az oldalakról elérhetővé tesz, vagy az oldalakba beépít adatbázisokból származó információkat. Ha az adatbázis jól felépített, a lekérdezés múltbéli időpontokra is működik. Az adatbázisok archiválásáról az előadás második részében beszélek. Az adatbázis, vagy CMS rendszer egy speciális formája a repozitórium - ezekről is a második részben kerül szó. A digitális könyv formában rögzített, archiválható információ is bekerülhet a repozitóriumba, és ezúton visszajuthat a webre, a hálózaton is kereshető és elérhető lehet.

B.) A mély web archiválása - adatbázisok, repozitóriumok

Az adatbázisok és a repozitóriumok információinak hosszú távú megőrzésére sem a webarchiválás kínálja a legjobb megoldást. Léteznek módszerek a relációs adatbázisok archiválására, a repozitóriumok pedig számos belső és külső archiválási lehetőséget támogatnak.

Itt kiindulási pontként leszögezhetjük, hogy a jelentős adatbázisokat, repozitóriumokat működtető intézmények feltehetően az átlagosnál valamennyivel archiválás-tudatosabbak, illetve az adatbázis vagy különösen a repozitórium szoftverek technikailag jobban támogathatják az archiválást, mint az egyszerű weboldalak.

Alapfogalmak

- Sötét archívum

Az adattartalom folyamatos, a külvilág számára nem látható archiválására szolgál. Az adatgazda végzi, de külső archiváló félhez kerülnek a tartalmak

- Végrendelet

Szabályozás az archívum megszűnésének esetére – mi legyen az adatok sorsa? A sötét archívummal együtt alkalmazandó

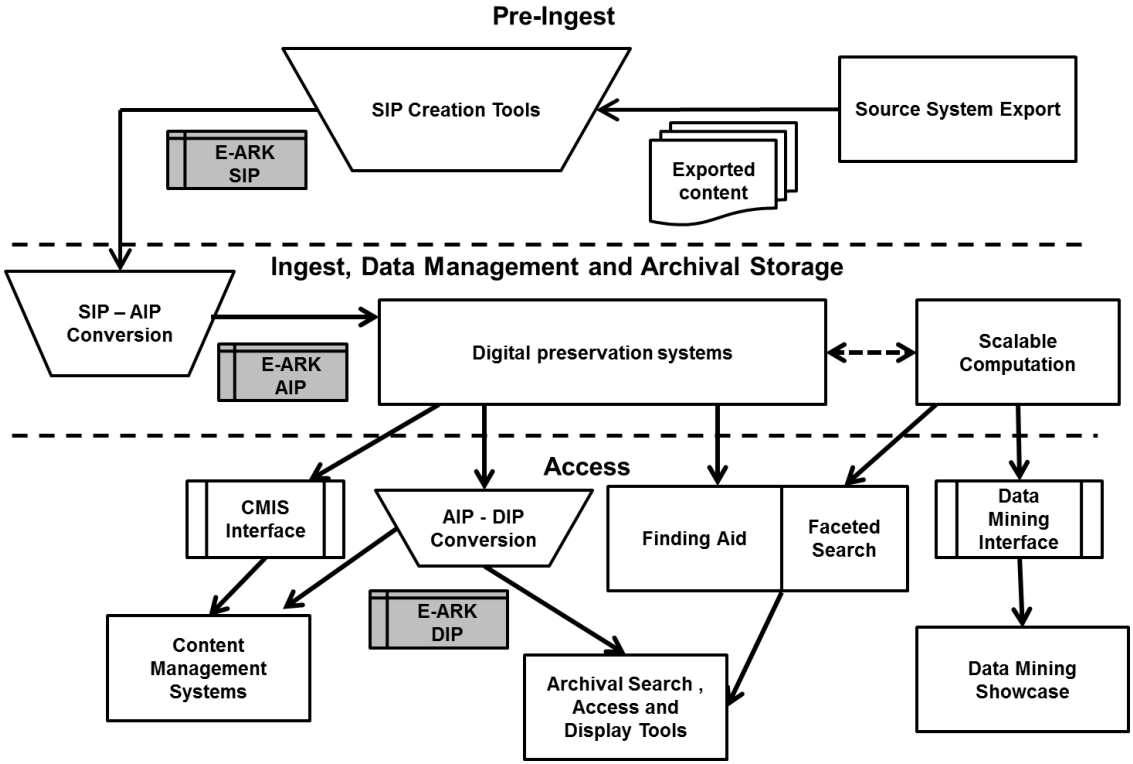
- Sírkő

Rövid információ hosszú távú fenntartása az adatbázis megszűnt weblapján – hol találhatóak a megőrzött információk?

- Adatbázis archiválás "belülről"

Az adatbázist üzemeltető intézmény tudatosan törekedhet a tárolt információk archiválására, a dinamikusan változó adatbázisok korábbi állapotainak megőrzésére. Lux (2018) relációs adatbázisok archiválására alkalmas technológiát ismertet.

Implementation of new technologies to ensure the sustainability of digital content, Lux Z., CDA2018, Bratislava, 2018. ISBN 978-80-89303-67-0, Univerzitna Kniznica.



E-ARK

Id. Lux Z.
MNL

- Repozitórium archiválása “belülről”

Azon túl, hogy a repozitóriumok különböző informatikai mentési eljárásokon keresztül, amelyeket történhetnek az adott szoftver tevőleges közreműködésével és a nélkül, képesek meghibásodásokat és katasztrófákat "túlélni", valódi archiválási lehetőségeket is támogathatnak. Olyan lehetőségeket, amelyekkel egy későbbi időpontban a tartalom (adatok és metaadatok) visszaállíthatóak még akkor is, ha a szoftverplatform már nem áll rendelkezésre.

Az MTA KIK-ben alkalmazott EPrints a tárolt tételeket - metaadatokat és a digitális dokumentumokat - képes XML formátumba exportálni (a dokumentumok karakteres formába átkódolva kerülnek az XML fájlba). Ha ezeket a szoftverfüggetlen mentéseket megfelelő helyen - akár másik félnél - megőrizzük, a tartalom helyreállítható lesz egy másik repozitóriumban.

```
-<eprints>
 -<eprint id="http://real-j.mtak.hu/id/eprint/13934">
    <eprintid>13934</eprintid>
    <rev_number>1</rev_number>
  -<documents>
   -<document id="http://real-j.mtak.hu/id/document/107801">
      <docid>107801</docid>
      <rev_number>1</rev_number>
    -<files>
     -<file id="http://real-j.mtak.hu/id/file/150320">
        <fileid>150320</fileid>
        <datasetid>document</datasetid>
        <objectid>107801</objectid>
        <filename>APH-2011-1.pdf</filename>
        <mime_type>application/pdf</mime_type>
        <hash>100972f55cc2a62854bf9cb9c4e675b7</hash>
        <hash_type>MD5</hash_type>
        <filesize>2090699</filesize>
        <mtime>2019-11-12 10:02:16</mtime>
        <url>http://real-j.mtak.hu/13934/1/APH-2011-1.pdf</url>
      -<data encoding="base64">
         JVBERi0xLjYNJeLjz9MNCjUyNSAwIG9iag08PC9MaW5lYXJpemVkIDEvTCAyMDkwNjk5L08gNTI3
         L0UgNTc4NDA2L04gNTIvVCAyMDg5NTkzL0ggWyA0ODUgMTMyOV0+Pg1lbmRvYmoNICAgICAgICAg
         IA01MzcgMCBvYmoNPDwvRGVjb2RlUGFybXM8PC9Db2x1bW5zIDUvUHJlZGljdG9yIDEyPj4vRmls
         dGVyL0ZsYXRlRGVjb2RlL0lEWzxFQjI4NkIxNkUxMEVFREUyRDlFRTU5RjhDRjQ4NkU5RD48MEQx
         OEZENTczNDMzNDQ3MDg3MkQ1QTA3MDM3MDkwOTA+XS9JbmRleFs1MjUgMjhdL0luZm8gNTI0IDAg
```

- Repozitórium archiválás "kívülről"

A repozitóriumok metainformációi az OAI-PMH protokollon keresztül arathatóak. Maguk a teljes szövegű dokumentumok is begyűjthetőek egy aggregátor által. Van de Sompel (2016) az OAI-PMH-n túl további lehetőségeket is ismertet.

Web Infrastructure to Support e-Journal Preservation (and More), Herbert Van de Sompel, David S. H. Rosenthal, Michael L. Nelson.
https://arxiv.org/abs/1605.06154

A repozitóriumok külső archiválása (és kollektív, kölcsönös archiválása) már régen megoldott a CLOCKSS és a LOCKSS kezdeményezések segítségével.

CLOCKSS: A Trusted Community-Governed Archive
http://www.clockss.org/

CLOCKSS

LOCKSS: Lots Of Copies Keep Stuff Safe http://lockss.org/

Lots of Copies
Keep Stuff Safe

A hazai dokumentum adatbázisok esetében bár léteznek tartalmi átfedések, alkalmanként jelentős lehet redundancia, az archiválás még nem megoldott. A teljes szövegek összegyűjtésére - egyelőre nem archiválás, csupán a kereshetőség biztosítása céljából - történt kezdeményezés.