

Saját készítésű tematikus "internet konzervek" a tanórán

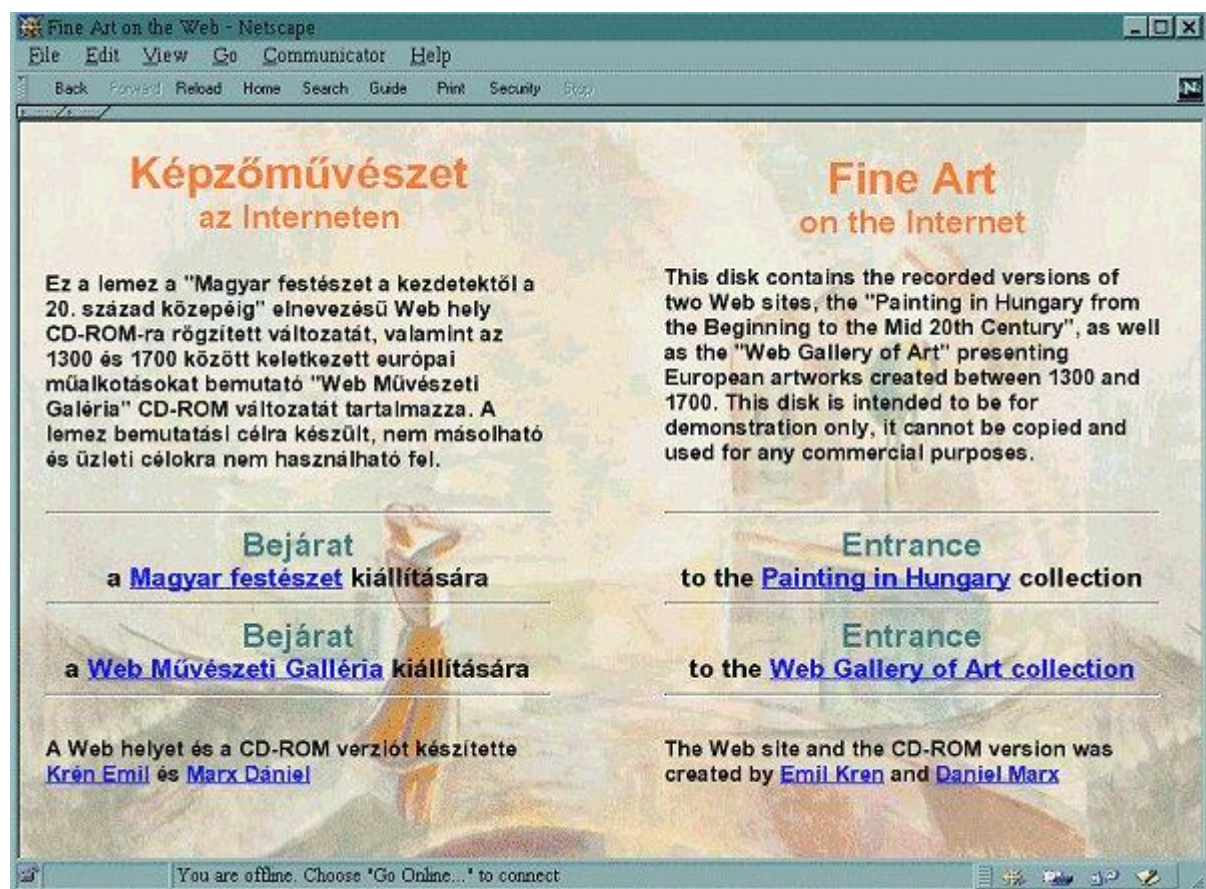
Dr. Kokas Károly

kokas@bibl.u-szeged.hu

JATE Egyetemi Könyvtár, Szeged * <http://www.bibl.u-szeged.hu>

(Az 1998. március 30-31-én megrendezésre került "SULINET – ablak a világra" konferencián elhangzott előadás írott változata.)

Az alábbi rövid írásban vázlatosan annak módszertanát próbálom elmondani, hogy miért és hogyan készítsünk saját Internet-anyaggyűjtéseket, azokat hogyan kezeljük, hogyan írjuk fel CD-re ill. hogyan szolgáltatassuk hálózaton. Mivel munkahelyemen (és általában az egyetemi világban) unixos rendszerekkel dolgozunk, a kialakult módszertan leginkább erre támaszkodik. Merthogy azonban a Sulinet világában nem ez, hanem az MS operációs rendszerei az uralkodók, igyekszem a Windows, Windows 3.11 ill. 95 vagy az NT problémáira is kitékinteni. Természetesen nem hallgatható el az a tény, hogy a vegyes operációs rendszerek ill. a Microsoft azon politikája, hogy nem föltétlen követi a kialakult szabványokat és saját formátumok tömegét használja, szóval mindez nehezíti dolgunkat. Azonban ma már rendelkezésünkre áll a középiskolákban is egy gyakorlatilag ingyenes eszköz, a Linux, amellyel meglévő hardver parkunkkal is kipróbálhatjuk a unixos megoldásokat.



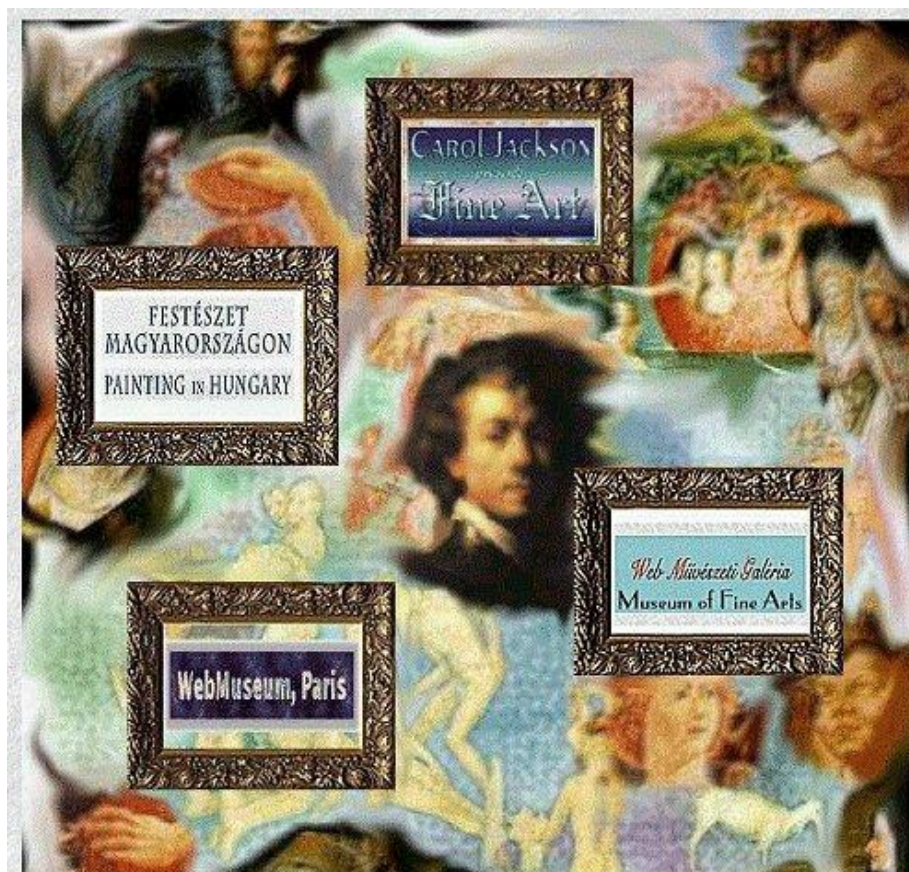
A KFKI szerverén található [két webszolgáltatás](#) CD-verzióját az alkotók készítették

1. Miért kell "konzerválni"?

Ma még nagyon nehéz egy tanórát úgy elképzelni, hogy a tanulók valós időben Interneten oldják meg a feladatokat. Erre az Internet-csatlakozással felszerelt kabinetben végzett metódusra igazából csak akkor van szükség, ha kifejezetten magát az Internetet tanítjuk, mondjuk számítástechnika órán. Jóval izgalmasabb azonban szakmailag, amikor az internetes információforrásokat a videóhoz, a diához, fóliához hasonlóan egyszerűen csak be akarjuk építeni az órai ismeretanyagba, akárcsak pusztán illusztrációként. Nyilvánvaló ez esetben szükséges és elégséges feltétel az, hogy a tanári gép képe legyen jó minőségben kivetíthető. Jó, ha ez a gép élő internetes kapcsolattal is rendelkezik, de az sem tragikus, ha nem.

A tanóra kiszámított, pergő menetében amúgy is nehezen fér be egy sok bizonytalansággal járó "valódi" Internet-keresés. A diákok sem erre kíváncsiak ez esetben, látják, tanulják ezt a megfelelő tanórán. Egy történelem vagy biológia órán a szaktanár – aki nem föltétlen komputer specialista –, nagyobb biztonságban érzi magát, ha előre letöltött, az iskola szerverén előkészített, ne adj Isten, megszerkesztett anyagot használhat.

2. Egy klasszikus módszer: a tükrözés



A *mirroring* vagy tükrözés régóta ismert módszer. Legfontosabb funkciója, hogy a leterhelt vonalakat kímélendő ill. a kockázat nélküli rendelkezésre állást biztosítandó egy nagyon gyakran használt információforrást vagy annak egy részéről "tükörképszerű" másolatot készítünk. Ez a másolat lehet hivatalos is (*official mirror site*) vagy magánhasználatú. Ez utóbbi esetben nem feltétlen kell engedélyt kérnünk, de nem is szolgáltatjuk tovább a hálózaton kifelé az adott anyagot.

A legegyszerűbb eset tehát, amikor a tanórán vagy a tanítási folyamatban később

felhasználható anyagot az iskola szerverére tükrözzük, hogy a későbbi zavartalan használatot biztosítsuk. Ez a "használat" legtöbbször további feldolgozást jelent, hiszen a letükörözött site legritkább esetben alkalmas közvetlen felhasználásra.

Nagyon sokféle tükröző program ismert az Internet világában, az egyszerű perl scriptektől a szép grafikus felületekkel és legördülő menüvel felszerelt csodáig bezárólag. Az egyik legismertebb és méltán

legnépszerűbb a WGET.

2.1 A WGET példája

A korábban GETURL néven ismert WGET Hrvoje Niksic <hniksic@srce.hr> horvát programozó munkája, tipikus GNU licence alá eső program. Így természetesen forrásnyelven is hozzáférhető. A grafikus felülethez szokott Windows-felhasználónak meglepő lehet a segédprogram mérete (100 kb-ot alatti!) és az is, hogy egyszerű parancsnyelvi interfésszel látták csak el. A WGET a különféle unixos szervereken használható elsősorban, gyakorlatilag bármelyik Unix verzió alatt lefordítható és portolható. Így az igen népszerű és elterjedt Linux különféle fajtái alatt is.

A program gyakorlatilag bárhol föllelhető az Interneten, mégis adjuk meg "hivatalos" site-ját:

<[URL:ftp://gnjilux.cc.fer.hr/pub/unix/util/wget/](ftp://gnjilux.cc.fer.hr/pub/unix/util/wget/)>

ill. egy tükrözés:

<[URL:ftp://sunsite.auc.dk/pub/infosystems/wget/](ftp://sunsite.auc.dk/pub/infosystems/wget/)>

<[URL:http://sunsite.auc.dk/ftp/pub/infosystems/wget/](http://sunsite.auc.dk/ftp/pub/infosystems/wget/)>

A programmal kapcsolatos tudnivalókkal, a fejlesztéssel külön levelezőlista is foglalkozik, a <wget@sunsite.auc.dk> címen, amelyre <wget-request@sunsite.auc.dk> címen kell feliratkozni.

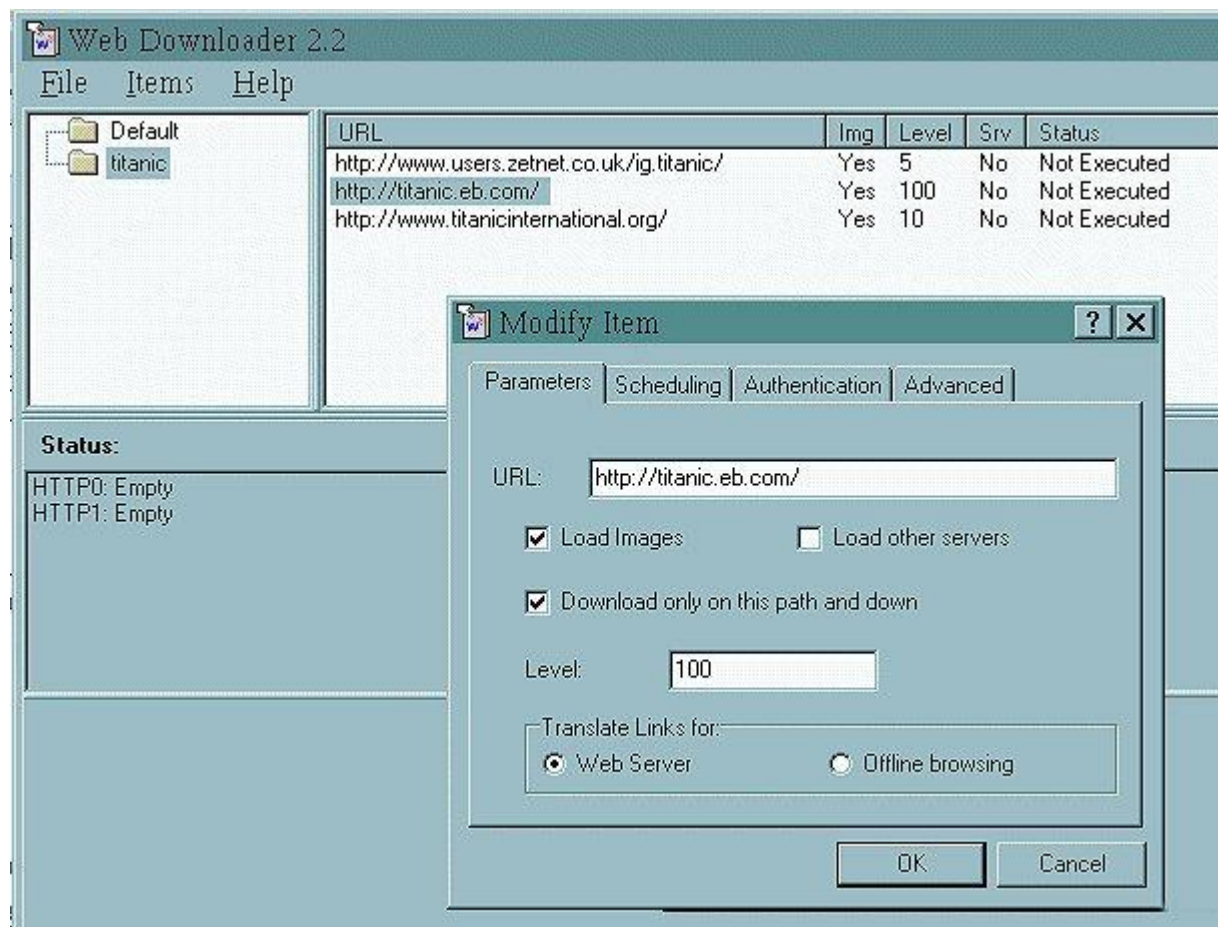
Mit tud tehát a WGET? Nagyon sokat és azt nagyon biztosan. A WGET egyszerű kezelési felülete ellenére maradéktalanul teljesíti mindazt, amire http ill. ftp site részleges vagy egészben való tükrözése során szükségünk lehet. Indíthatjuk természetesen batch file-ból, így időzítve is. Nagy előnye, hogy az alfanumerikus felület miatt pl. akár otthonról dialup hívással is elindítható a program. Paramétereit nagyon sok kombinációban, a szokásos unixos flagekkel adhatjuk meg. Megadhatjuk pl., hogy egyszerű letöltés vagy tükrözés legyen-e a munkafolyamat eredménye, hogy csak az adott könyvtártól lefelé menjen ill. eltérjen-e az illető site-tól. Milyen file-típusokat hozzon ill. ne hozzon le, konvertálja-e relatívvá vagy abszolúttá a linkeket stb. stb. A program nagyon alkalmas tükrözött helyek karbantartására, ezt elősegíti nagyon részletesen felépített logfile-ja is.

2.2 Web-letöltés Windows alatt

Persze a windows-os világban is találunk hasonló célból megírt programokat, még ha kompaktságuk, kifinomult beállítási lehetőségeik nem is mérhetők a WGET-éhez. Ha megnézünk csak két ismert "shareware lelőhelyet", mint pl. a TUCOWS-ot (van magyar tükör! : <http://www.euroweb.hu/tucows/>) és a Windows95 site-ot (<http://www.windows95.com>) számtalan ajánlattal találkozunk, általában az "off line

browser tools” terminológia alatt.

A *WebZip 2.0* ravasz módon azzal hidalja át az eddigi problémákat, hogy a letöltött anyag egy nagy zipfile-ba, mint egy konténerbe kerül, így aztán azt már nyugodtan felírhatjuk CD-ROM-ra. A *WebMirror 1.15* “klasszikus” tükröző program, de nem túl sok beállítási lehetőséggel, hasonlóak ehhez a *WebVCR* és a *Grab-a-Site* is. A *WebWhacker* már kifinomultabb program, sőt több mint négy megabájtos méretével egyenesen tekintélyt parancsoló. Ez már közelebb van a WGET tudásához. A *NetAttaché* programnak a shareware használati idő lejárta után külön “litte” verziója is van, ami freeware-nek számít. Régióta kedvelt profi program a web-letöltők között a *TelePort Pro*, amivel szinte mindent meg lehet csinálni. S a sort folytathatnánk, hiszen a két említett site-on legalább 30 hasonló program van.



A Web Downloader 2.2 Windows 95/NT alól a legjobb választás

Mégis szembetűnik egy, ami már csak freeware jellege miatt is figyelemreméltó. A **WEB DOWNLOADER 2.2** kompakt méretével (200 kilobájtos lévén, akár floppyról is indítható) és ügyes kezelőfelületével hívja föl magára a figyelmet. Beállíthatjuk a tükrözendő helyet itt is, megadhatjuk, hogy más szerverre átmenjen-e ill. az adott könyvtárnál “följebb” szabad-e mennie a letöltésnek. Indíthatjuk időzítve is a programot, sőt a rendszergazdák és “más hálóőrök” nagy bánatára elvileg 100 csatornán egyszerre is megindíthatjuk a tükrözést. A html lapokba ágyazott dokumentumtípusok itt is elkülöníthetőek.

3. Saját anyagot szerkesztünk

Az alábbiakban röviden egy művészettörténeti példa kapcsán szeretném bemutatni a munkamódszert.

A művészetek története elég általános téma ahhoz, hogy sokak számára hasznosítható anyagot készíthessünk, s az is vonzó volt, hogy számtalan nagyon jó Internet szolgáltatást találtunk ebben a témakörben. Az első lépés az volt, hogy könyvtárunk "hivatalos" magyarországi tükörszervere lett két ragyogó gyűjteménynek, a *Carol Jackson-féle "Fine Art"* anyagnak és az ún. párizsi "*Webmuseum*" kollekciónak. Mivel hazánkban is van két ezekhez mérhető nagyon szép hely, Krén Emil fizikus által az NKA segítségével szerkesztett magyar festészettörténeti tárlat és ennek európai megfelelője, úgy döntöttünk a két rendszeresen tükrözött helyhez és e két utóbbi magyarhoz közös "[bejáratot](#)" készítünk.



A későbbiek során jött az ötlet, hogy kiegészítsük e gyűjteményeket, más kisebb virtuális múzeumok kollekciónak, úgy, hogy az adjon egy 5–6000 képet tartalmazó, minden lehetséges szempontból "alapképtárnak" tekinthető saját anyagot. Végül is e gyűjtő és szerkesztő munka eredményeképpen egy majdnem 700 megabájtos "múzeumot" kaptunk. Igazából hét nagyobb összeszerkesztett képtárról van szó, amelyek közül említettem már a *Carol Jackson-félet*, a

WebMuseum-ot. Ezekhez jöttek még a *Louvre*, a *Vermeer Múzeum* ill. az *Uffizi* képei és két összeállítás *Leonardóról* és *Monet-ről*.

Az már egyéni megfontolás kérdése persze, hogy mennyire szerkesszük kompakttá az anyagunkat. Összefésüljük-e a párhuzamosságokat stb. Magam azon a véleményen vagyok, hogy szerencsésebb, ha ennél inkább megmarad az eredeti hely struktúrája. Így inkább arra szorítkozunk, hogy ellenőrizzük, mennyire hibátlan az anyag (az "üres" linkek kikeresésére rengeteg jó shareware és freeware program van a hálózaton) és szerkesszünk az együttes használatot biztosító (ha szükséges akár többféle) tartalomjegyzéket.

4. CD-ROM-ra írjuk a gyűjteményt

Mint említettem az összegyűlt anyag 700 megabájt körüli volt, így az adott szerver helyproblémáit is figyelembe véve (hiszen az eredeti tükrözéseket közben is megtartottuk) jó megoldásnak látszott, hogy az egész kollekciónak CD-ROM-ra írjuk. Mivel az volt a cél, hogy ezt esetleg később gond nélkül vissza tudjuk másolni a szerver winchesterére ill. más módon szolgáltatni tudjuk valamelyik unixos szerverünk segítségével, ezért a Unix file-struktúrával kompatibilis és amúgy is szabványos *Rockridge CD formátum* mellett döntöttünk

(Az adott példában szereplő anyag unix alatti előkészítésében és CD-re írásában Zawiasa Róbert kollégámmal dolgoztunk együtt.)

4.1 Formátumproblémák

A CD-ROM formátumgondok onnan adódnak, hogy az Interneten található anyagok – unixos szolgáltatójuk lévén – magától értetődően ismerik és használják a DOS és a Windows világában ismeretlen hosszú file- és alkönyvtárneveket és persze a kis- és nagybetűt megkülönböztető írásmódot. Az eredeti ISO 9660 szabványú CD-főírás szabvány ezt nem ismerte, ragaszkodott a 8+3 formulához. A hosszú és vegyes írásmódú filerendszer ugyan hamarosan CD-re írható lett a Rockridge kiterjesztés révén. Ezt a szabványt azonban nem használják a Microsoft operációs rendszerek, hanem saját kiterjesztésekkel kísérleteznek, így a *Joliet* ill. *Romeo* formátumokkal. A gyakorlatban ez azt jelenti, hogy minden rendszer számára jól olvasható (hosszú filenevekkel is operálni tudó) filerendszer ma nem létezik CD-én. Az összefüggések hosszas magyarázata helyett nézzünk meg egy táblázatot:

Filerendszer	DOS	Windows 3.11	Windows 95	Windows NT 3.51	Windows NT 4.0	Mac	Unix
ISO 9660 8+3	Igen	Igen	Igen	Igen	Igen	Igen	Igen
Más MS-DOS 8+3	Igen	Igen	Igen	Igen	Igen	Igen	Igen
Joliet (Adaptec Toast, Easy CD Creator, Easy CD Pro 95)	Igen, rövid névvel	Igen, rövid névvel	Igen	Nem	Igen	Igen, rövid névvel	Igen, rövid névvel
Joliet (CD Creator 2.x)	Nem	Nem	Igen	Nem	Igen	?	?
ún. long file name mód (Easy CD Creator)	Nem	Nem	Igen	?	Igen	?	?
Romeo (Easy CD Pro 95)	Nem	Nem	Igen	Igen	Igen	Igen, ha 31 karakternél rövidebb	vegyesen látszik

Macintosh HFS (Adaptec Toast)	Nem	Nem	Nem	Nem	Nem	Igen	Nem
UDF (DirectCD)	Nem	Nem	Igen, spec. driverrel	Nem	Igen, spec. driverrel	Igen, spec. driverrel	?
ISO 9660 Level 3 (DirectCD for Windows)	Nem	Nem	Igen	?	Igen	Nem	?
Rockridge (mint Unix alatt legenerált image felírva)	Igen, rövid névvel	Igen, rövid névvel	Igen, rövid névvel	Igen, rövid névvel	Igen, rövid névvel	?	Igen

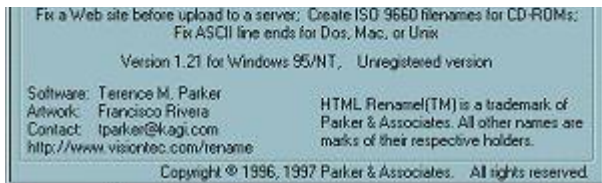
(A táblázat a <http://www.adaptec.com/tools/compatibility/> site adatainak felhasználásával készült. Sajnos az OS/2-ről nincsenek közelebbi információim, de a Rockridge formátumot a hírek szerint nem ismeri.)

4.2 Mi a megoldás?

Mint az jól látható, sajnos igazi korrekt megoldás jelenleg nincsen. Szögezzük le előljáróba, amikor már hozzá kell nyúlni az anyaghoz név- és filestruktúra szinten, az sok bajnak lehet okozója. Viszont elképzelhető, hogy másképpen nem boldogulunk. Ha teljesen biztonsággal akarunk eljárni, az egész anyagot át kellene dolgoznunk kézzel 8+3-as formátumúvá, hogy ti. "azzal biztos nem lesz baj". De képzeljük ezt el 8-10.000 file esetében! S akkor még a html-be ágyazott linkek kusza szövedékéről nem is volt szó.



Erre a gondra is találunk azonban a világhálón egy célprogramot, a shareware disztribúcióval terjesztett **"HTML Rename!"** nevűt. Ez a 32 bites Windows-os program nem csupán a file-ok átnevezését, automatikus lerövidítését végzi el, hanem közben ennek megfelelően átalakítja a html lapokban a hivatkozásokat is. Más hasonló programok közül ezt találtam legbiztosabban működőnek. Különleges tulajdonsága még, hogy olyan speciális kérdéseket is megoldhatunk vele "útközben", mint pl. a Mac vagy 32



bites Windows filestruktúrának megfelelő átnevezések.

Természetesen tudomásul kell vennünk, hogy a HTML Rename! "végsőlegesen" átalakítja anyagunkat, ezért a művelet teljes befejeztéig ne töröljük az eredeti

tükrözött vagy letöltött anyagot

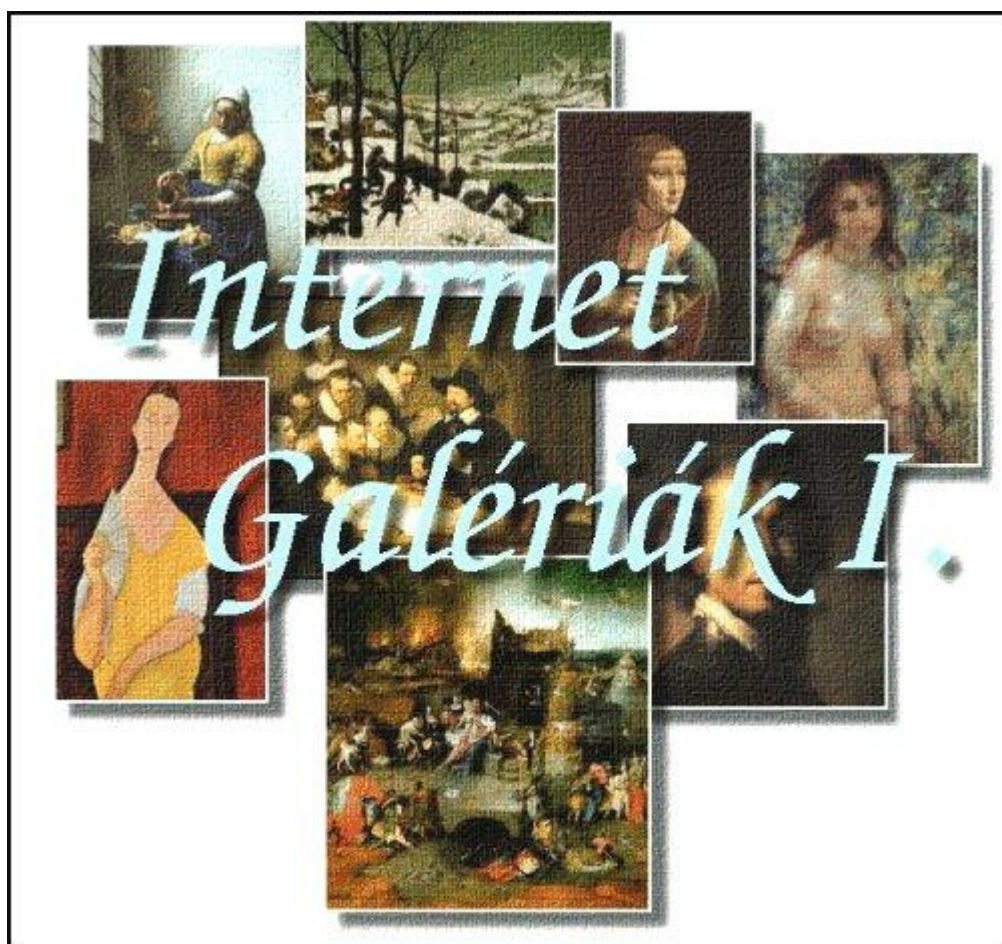
6. Szolgáltatás a hálózaton

Elkészült anyagunkat általában csak magánhasználatra szánjuk, hiszen nem valószínű, hogy az összes forráshelyről engedélyt kértünk, s a belső iskolai használat miatt erre szükség sem nagyon van.

Mindenesetre, ha nyilvánosság számára készítjük a szolgáltatást, ne feledkezzünk el a hálózaton is létező szerzői jogi kérdésekről, s tanulmányi/iskolai használatra hivatkozva kérjünk ingyenes hozzájárulást, amit a szerverünkön jól látható helyen és a források megjelölése mellett helyezünk el!

Ha a nyilvánosság számára nem is, de felmerülhet az igény, hogy az iskolai intranet hálózatban mások számára is elérhetővé tegyük az anyagot. Windows NT szerver esetében a winchesteren tárolt anyag, vagy a megfelelő NT verzió alá felírt CD különösebb nehézség nélkül elérhető, ha vállaljuk az átírás és javítás igen munkaigényes és kockázatos módját (ami meg a célszoftver alkalmazásával sem föltétlen egyszerű).

Érdekes itt röviden még arra is kitérni, hogy az ilyen nagy terjedelmű anyagok összeszerkesztése igen nagy figyelmet igényel, s kézi munkával nagyon nehezen megoldható. Ezért érdemes azokat a világhálón (az említett shareware helyeken pl.) fellelhető ellenőrző szoftvereket használnunk, amelyek egyenként kontrollálják a hibás linkeket és kigyűjtik a hiányzó képek és hivatkozások listáját stb.



A Szegedi Egyetemi Könyvtárban elkészült első "internet konzerv" címlapja.

Jóval egyszerűbb és az eredeti struktúrákat legjobban megőrző mód azonban, ha az elkészült anyagot az említett Rockridge formátumban írjuk fel CD-re, és így szolgáltatjuk. Ennek elkészítésre és szolgáltatására is kiválóan alkalmas a Linux. Ezzel egyrészt könnyedén legyárthatunk egy a Rockridge szabványú CD kiterjesztésnek megfelelő image file-t, amelyet aztán bármelyik CD-ROM író programmal felírhatunk egy nyers CD-re. Az így elkészült CD-t egyszerű mountolással részévé tehetjük a linuxos webszerverünk file-rendszerének, így az – pl. egy szimbolikus linkkel – bárhova beilleszthető a saját web-struktúrába. A mai nagy sebességű CD-ROM meghajtók lehetővé teszik, hogy viszonylag nagyszámú kliens egyszerre érje el az adott CD tartalmát.

Az már az elérhetőség és a nyilvános link kérdése, hogy adunk-e lehetőséget arra, hogy kívülről, az egész Internet közösség számára láthatóvá tegyünk művünket, vagy megelégszünk egy "láthatatlan", csak a szűkebb közösség által ismert URL megadásával.

Szeged, 1998. április 13.

