

AZ NJSZT INFORMATIKATÖRTÉNETI FÓRUM HONLAPJÁNAK MEGŐRZÉSE AZ „ÖRÖKKÉVALÓSÁG” SZÁMÁRA

Dömölki Bálint†

Havass Miklós

Langer Tamás

Neumann János Számítógép-tudományi Társaság, Informatikatörténeti Fórum

Holl András

Magyar Tudományos Akadémia Könyvtár és Információs Központ

Absztrakt

A Neumann János Számítógép-tudományi Társaság Informatikatörténeti Fóruma (iTf) adatokat gyűjt egy Adattárba (iTÁ) a magyar informatikatörténet tényeiről. A gyűjtött adatokat egy WordPress rendszeren alapuló honlapon mutatja be napjaink érdeklődőinek. A több mint 15 éves gyűjtőmunka során sok ezer adat került az adattárba, a hazai informatika történetében szerepet játszó személyekről, intézményekről, termékekről, a lebonyolított rendezvényekről, valamint a keletkezett írásokról és videókról. Felmerült bennünk a kérdés, hogy számolva a használt eszközök, intézményi formák változásával, mit kell tenni azért, hogy az információk minél tovább megmaradjanak, elérhetők legyenek az utókor számára. A probléma nyomán közös projektet („örökkévalóság” projekt néven) indított az iTf a Magyar Tudományos Akadémia Könyvtár és Információs Központtal (MTAK), amely során az iTÁ teljes tartalma megfelelő formában elhelyezésre került az MTAK repozitóriumban. Előadásunk ennek a megoldásnak az elveit, „almanach”-okon alapuló technológiáját és egyes lényegi lépéseit ismerteti, valamint utalunk a megőrzés további lehetőségeire.

Abstract

Preserving the Website of the IT History Forum (iTf) of the John von Neumann Computer Society “for Eternity”

The IT History Forum (iTf) of the John von Neumann Computer Society collects data for a data archive (iTÁ) on the facts of Hungarian IT history. The collected data is presented to today's interested audiences on a website based on a WordPress system. Over more than 15 years of data collection, several thousand entries have been gathered in the repository – about individuals, institutions, and products that played a role in the history of Hungarian computing, as well as events, written works, and videos produced over time. This long-term effort raised the question: given the changes in tools and institutional frameworks, what

must be done to ensure that this information remains accessible and preserved for future generations. As a response to this challenge, the iTF launched a joint project – named the “Eternity Project” – in collaboration with the Library and Information Centre of the Hungarian Academy of Sciences (MTAK). As part of this project, the entire content of the iTA was transferred, in an appropriate format, to MTAK’s repository. Our presentation outlines the principles behind this solution, the technology based on “almanacs,” and the key steps involved. We also touch upon further possibilities for long-term preservation.

Előzmény

A magyar számítástechnikusok 1968-ban alapították meg első tudományos egyesületüket, a Neumann János Számítógép-tudományi Társaságot (NJSZT). A Társaság célja, hogy intézményektől független szakmai fórumként segítse hazánkban, illetve a magyar nyelvterületeken a számítástechnikai, informatikai kultúra elterjedését, lehetőséget biztosítson a hazai számítástechnikusok közötti eszmecserék számára. A Társaság szakosztályokban, területi szervezetekben fejti ki tevékenységét.

Néhányan a Társaság tagjai közül önálló szakosztályként 2009 januárjában létrehozták az Informatikatörténeti Fórumot (iTF), abból a célból, hogy tagjai életpasztalataik, emlékeik segítségével gyűjtsék, rendszerezék és bemutassák a magyar informatikatörténet eseményeit. A Fórum kiemelt feladatának tekinti, hogy bemutassa a múltban fontos szerepet játszó személyeket, intézményeket, termékeket, rendezvényeket és írásokat. A „rovatokba” gyűjtött és rendszerezett információkat a Fórum digitális Adattárba (iTA) szervezi, és egy WordPress rendszeren alapuló honlapon (itf.njszt.hu) mutatja be. Az Adattárban az információk Adatlapok formájában a következő Rovatokba vannak rendezve: Személyek, Intézmények, Termékek, Írások, Rendezvények, Videótár.

Az információk gyűjtése mellett a Fórum emlékkonferenciákon mutatja be a magyar számítástechnika jelentősebb műhelyeit, fontosabb szakterületeinek történetét, az előadások videófelvevételeinek és prezentációinak publikálásával (iTF Események), és videóriportokat készít kiemelkedő, meghatározó teljesítményt nyújtó szakemberekkel (Arcképek a magyar informatika történetéből).



1. ábra: Az iTF honlapjának szerkezete

A történeti adatok és információk bemutatása céljából – amíg az NJSZT és azon belül az iTF fennmarad – az iTF üzemelteti a fent említett honlapot. A honlap szerkezetét az 1. ábra szemlélteti. Korunk gyors ütemű változásai Társaságunkat is átalakulásra (például egyesülés, osztódás, megszűnés) készíthetik. Az Adattárban összegyűjtött és feldolgozott adatok hosszú távú megőrzését azonban az iTF szerkesztősége függetleníteni szeretné a Társaság élettartamától, ezért azokat olyan szervezetre szeretné bízni, amelynek hosszú távú fennmaradását önálló nemzeti létünk szavatolja. Ennek szellemében megállapodást kötött az MTA Könyvtár és Információs Központtal (a továbbiakban MTAK) abból a célból, hogy az iTA adatait a MTA repozitóriumába telepítsék, biztosítva a nemzetközi szabványok szerinti hosszú távú visszakeresés lehetőségét.

Az MTAK repozitóriuma, a REAL (Holl, 2013) egy jelenleg nyolc gyűjteményből álló, az EPrints szoftveren alapuló digitális könyvtári rendszer, mely túlnyomó részben nyílt hozzáférésű dokumentumokat szolgáltat teljes szöveggel. A repozitórium (a digitális könyvtári katalógussal együtt) a könyvtár alapszolgáltatása, amelyet hosszú távon fenn kíván tartani. A hosszú távú megőrzés eszköze a PDF/A dokumentumformátum alkalmazása. A szolgáltatott dokumentumok láthatóságát a REAL-szolgáltatásra mutató hivatkozások jelentős száma, illetve a nemzetközi és hazai repozitóriumi hálózatokba való integráltság biztosítja.

Az áttelepítés lépései

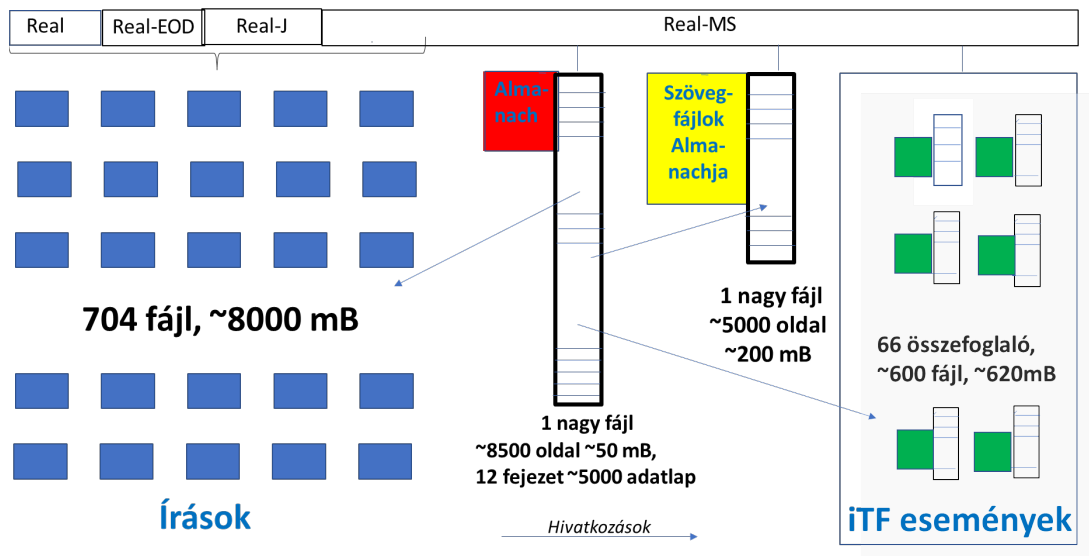
Az MTAK számára átadott fájlok az ott levő repozitóriumi gyűjtemények (REAL, REAL-MS, REAL-EOD, REAL-J) egyikébe kerülnek, és egy Leíró weboldal keletkezik róluk, amelynek tartalmából megismerhető a fájl MTAK-beli elérhetősége és az összes metaadata.

Informatikatörténeti munkánk az MTAK-ban a következő módokon jelenik meg:

1. Adattárunk Írások rovatának legtöbb eleme önálló tételként felkerült az MTAK repozitóriumába (például <https://real-ms.mtak.hu/29433/>).
2. Ugyancsak egyedenként, egy-egy fájlba tömörítve elhelyezésre kerültek a repozitóriumba az egyes Emlékkonferenciákra (iTF eseményekre) vonatkozó tartalmi információk (a program, az előadások prezentációi és a videófelvételekre történő YouTube-hivatkozások, például <https://real-ms.mtak.hu/29948/>),
3. Az Adatlapokon hivatkozott és a honlapunkon fellelhető egyéb kisebb dokumentumok (önéletrajzok, nekrológok, publikációs jegyzékek, rövidebb cikkek stb.) a Szövegfájlok Almanachjában kerültek összegyűjtésre. Ez az almanach is felkerült az MTAK repozitóriumába (<https://real-ms.mtak.hu/29253/>).
4. Adattárunk videótárának elemeit (a videóriportokat, az iTF-események videóit és az egyéb általunk összegyűjtött videókat) a YouTube-on tároljuk, amelytől azt reméljük, hogy üzemeltetője hosszú távon gondoskodik ezeknek az anyagoknak a folyamatos elérhetőségéről.

5. Valamennyi rovat valamennyi Adatlapjának a másolata a Magyar Informatikatörténeti Almanachba (a továbbiakban Almanachba) került összegyűjtésre, és egyben tárolásra az MTAK repozitóriumában (<https://real-ms.mtak.hu/30173/>). Annak érdekében azonban, hogy az Almanach a honlapunk létezése nélkül is használható legyen, az adatlap másolatokon valamennyi, az itf.njszt.hu médiatárjába mutató linket az MTAK környezetébe mutató URL-lé kell átalakítani. Ez a következő módon történik:
- A rovatok (személyek, intézmények, termékek, rendezvények, írások, videótár) egymás közti hivatkozásai az Almanachon belüli hivatkozásokkal vannak helyettesítve;
 - Az MTAK repozitóriumába felvitt írások (ld. 1. pont) hivatkozásai az MTAK-beli Leíró URL-jével vannak helyettesítve;
 - Az iTF eseményeken elhangzott előadásokra való hivatkozások az adott eseményt tömörítő fájl (ld. 2. pont) MTAK-beli Leírójának URL-jével vannak helyettesítve;
 - A Szövegfájlok Almanachjában összegyűjtött fájlokra (ld. 3. pont) való hivatkozások a Szövegfájl Almanach MTAK-beli leírójának URL-jével vannak helyettesítve. Az MTAK-beli Leíróban a könnyebb tájékozódás kedvéért egy táblázatban feltüntetjük, hogy az egyes szövegfájlok a Szövegfájlok Almanachjának hányadik oldalán kezdődnek. Például egy Kalmár Lászlóval kapcsolatos szövegfájltra történő hivatkozást a <https://real-ms.mtak.hu/29253/#page=2869> URL-lel helyettesítünk. Ezen keresztül eljutunk a Szövegfájl Almanach MTAK-beli leírójához. Innen letöltve a Szövegfájl Almanachot (download) megtudjuk annak az URL-jét, azaz jelen esetben <https://real-ms.mtak.hu/29253/12/SzovegfajlAlmanach.pdf>. Ehhez hozzátéve az oldalszámot eljutunk a megfelelő szöveghez: <https://real-ms.mtak.hu/29253/12/SzovegfajlAlmanach.pdf#page=2869>;
 - Az ezektől különböző, az itf.njszt.hu médiatárján kívülre mutató hivatkozások esetében helyettesítés nem történik. Ide tartoznak a 4. szerinti YouTube-linkek. Ezek a hivatkozások addig élnek, amíg az illető tartalom a az adott szolgáltatónál fellelhető.

Az Almanach és az általa hivatkozott kapcsolódó elemek elhelyezkedését az MTAK repozitóriumában a 2. ábra szemlélteti. Fontos megjegyezni, hogy a repozitóriumban elhelyezett Almanach feladata nem az iTF honlapjának helyettesítése, még ha az Almanachban lévő adatlapok meghívásával a honlapon lévő valamennyi adat megjeleníthető is. Ez azonban nem teljesen felhasználóbarát módon történik, de az Almanach célja a megőrzés az „örök-kévalóság” számára, és nem feltétlenül a felhasználói kényelem.



2. ábra: Az Adattár elemeinek elhelyezése a repozitóriumokban

Az MTAK-változat előállítás

Az Almanach létrehozásától szinte függetlenül folyamatosan történt és történik az írá-
sok felvitele az MTAK repozitóriumába. Ez, nagyobb csoportokat összeváruva, úgy zajlik,
hogy először a honlap Írások rovatának tartalmát exportáljuk egy Excel-táblába, majd ezt
átalakítjuk az MTAK által kíván formátumba, amiből az MTAK-ban működő algoritmus
automatikusan elvégzi az elhelyezést a megfelelő repozitóriumi gyűjteménybe, a táblából
kinyert fájlnev alapján csatolva magát a dokumentumot is. Ezután valamennyi kézimun-
kát igényel annak a táblázatnak az előállítás, amely összepárosítja az illető írás honlapbe-
li címét az MTAK-beli Leíró URL-jével.

A kihívást a PDF-formátumú Almanach és a Szövegfájlok Almanachjának létrehozása jelentet-
te. Az Almanach esetében össze kellett gyűjteni a több mint ötezer adatlapot egyetlen PDF-fájl-
ba és működőképpessé kellett tenni a PDF-fájlon belül az adatlapok közötti, valamint az írásokra,
az iTF-eseményekre és a szövegfájlokra az MTAK repozitóriumába mutató hivatkozásokat.

Hasonlóképpen össze kell gyűjteni egy Szövegfájlok Almanachja PDF-fájlba az összes, mint-
egy ezer darab szövegfájlt, és egy táblázatot létrehozni arról, hogy melyik szövegfájl hánya-
dik oldalon kezdődik. Ez utóbbira azért van szükség, hogy a megfelelő fájl megtalálását
a Szövegfájlok Almanachján belül segítsük.

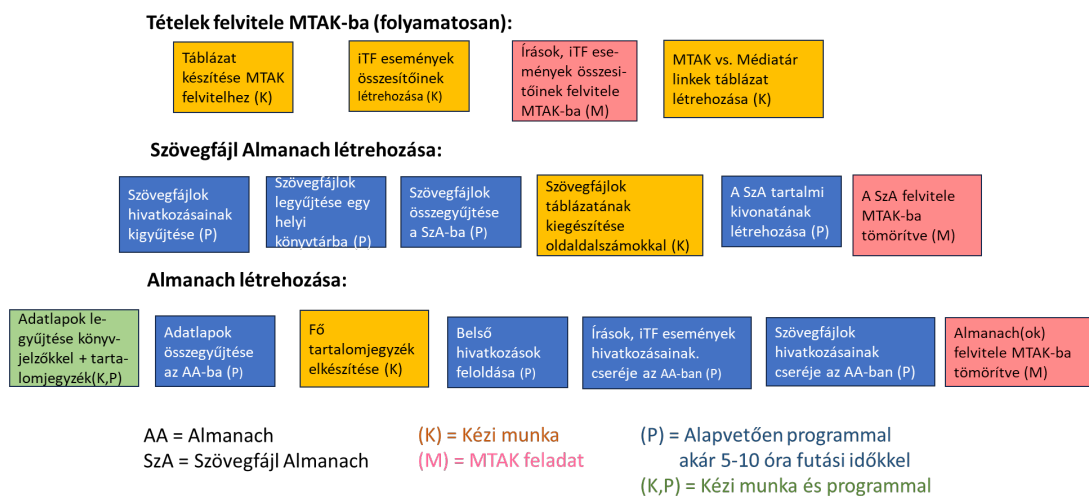
Feladat még az egyes iTF-eseményekhez tartozó összes információ összegyűjtése, azaz az
iTF-esemény adatlapja, kiegészítve a meglévő prezentációkkal az iTF-eseményt összefog-
laló PDF-fájlba úgy, hogy az adatlapból mintegy tartalomjegyzékből egy kattintással eljus-
sunk a megfelelő prezentációhoz.

A fentebb leírt feladatok automatizálható részét programmal, a nem automatizálható vagy nehezen automatizálható feladatokat kézi módszerrel oldottuk meg. Akkor is kézi módszerrel, ha az automatizálása több ráfordítást igényelt volna, mint kézzel végrehajtani (akár többször is) az adott lépést.

Az automatizált megoldás mintegy 2000-2500 sornyi Excel- és Word- (Visual Basic) makrót jelentett Word, Excel és Adobe Acrobat Pro használatával. Érdekesség: az egyes makrókat általában a ChatGPT-vel írtuk meg, néhány iteráció után elvezetve őt a megfelelően működő makró kódjához.

Az automatizált kilenc lépés az adathalmaz mérete miatt több órás-, esetenként akár teljes éjszakai futást jelentett, de a notebook-on a gépidő rendelkezésre állt, és a futtatás közben lehetett rajta mást is csinálni. A kézi megoldások (öt lépés) összességében jó néhány napot vettek igénybe. Egy új Almanach és Szövegfájlok Almanachja változat létrehozása mintegy két emberhétnyi munkát igényelt.

Az előállítás lépéseit a 3. ábra szemlélteti.



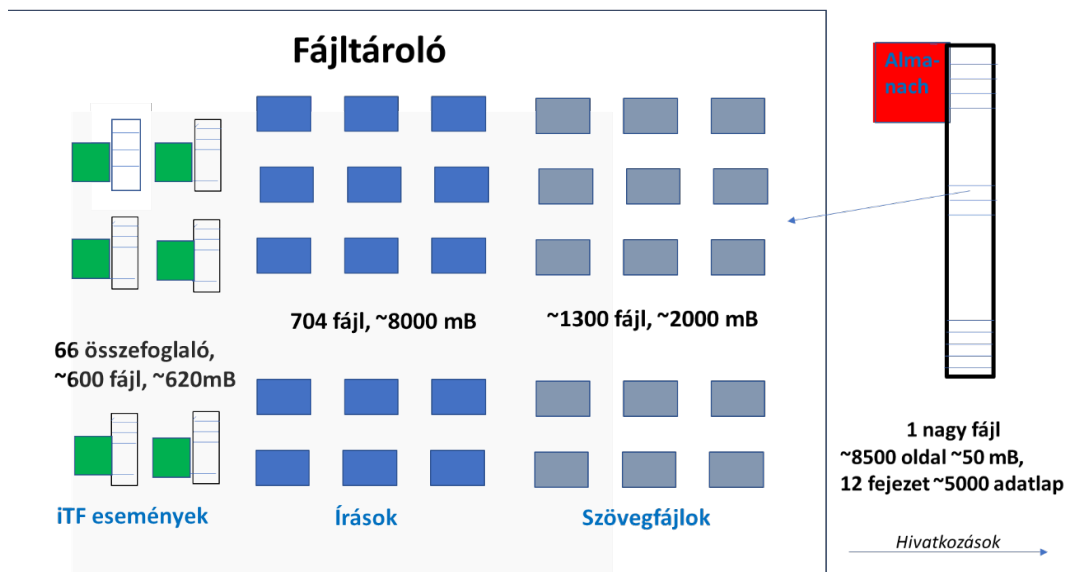
3. ábra: Előállítás lépései

Alternatív lehetőségek

Feltételezésünk, hogy ha véges energiával minél több megoldást találunk a honlapunk megmentésére az „örökkévalóság” számára, annál inkább nő az esélye, hogy utódaink valamelyikbe beleakadnak.

Alternatív lehetőség egy **Stand-alone** változat létrehozása. Egy csomagba összegyűjtjük magát az Almanachot, az írásokat, a szövegfájlokat, a prezentációkat és iTF-eseményeket egyenként külön-külön PDF-fájlokban. Az Almanachban a többi elemre való (nem belső)

hivatkozásokat a csomagban szereplő megfelelő fájl névre cseréljük. A Stand-alone változat különböző PDF-orientált archívumokba helyezhető el (például Zenodo, Figshare, Internet Archive). A Stand-alone változat szerkezetét a 4. ábra szemlélteti.



4. ábra: A Stand-alone változat szerkezete

Egy másik alternatív lehetőség a webhely statikus honlapként (HTML-oldalak gyűjteményeként), azaz a honlap oldalairól készülő pillanatképek halmazaként való előállítás. A statikus honlap mentésre és hosszú távú megőrzésre azért alkalmas, mert nem igényel háttérrendszert (például adatbázist vagy szerveroldali futtatást), így egyszerűen archiválható, elhelyezhető alapvetően nem archiválási célokra létrehozott különböző szervereken. A fenti alternatív lehetőségek esetében megfontolható valamilyen időtálló fizikai adathordozóra való elhelyezés és letétbe helyezés valamilyen erre kész tárolóhelyen.

A különböző megoldások összehasonlítását az 1. táblázat mutatja.

Megoldás	Előnye	Hátránya
MTAK-változat	Értékes elemek (írások, iTF-események) már napjainkban is használhatóak; az ügy iránt elkötelezett együttműködő magyar partner.	Kifejezetten egy könyvtárra szabott megoldás.
Stand-alone változat	Egy csomagban átvadható/eltehető archívumokba.	Értékes elemek (írások, iTF-események) külön fellelhetősége megszűnik.
Statikus honlap	U.a. mint a stand-alone változatnál + pontos lenyomata a dinamikus honlapnak.	U.a. mint a stand-alone változatnál + a PDF-orientált archívumok nem fogadják be.

1. táblázat

E két lehetséges megoldással kapcsolatban az iTF további vizsgálatokat tervez a jövőben.

Összefoglalás

A 2024. évben megvalósult az iTF-honlap az év nyári állapota szerinti tartalmának áttelepítése és tárolása az MTA repozitóriumában. A tartalmak megőrzése egy proaktív stratégia szerint történt, ahol a REAL mint archiválásra és szolgáltatásra szolgáló repozitórium, valamint a PDF/A médiaformátum használata lehetővé teszi az iTF tartalmainak hosszú távú megőrzését. A megőrzésen kívül a repozitóriumi tárolás haszna az egyes elemek jelen idejű megtalálhatóságának javításában is megnyilvánul (pl. Google Scholar-elérés).

A repozitóriumi tárolás előállításának fentiekben leírt folyamatát lehetőség szerint célszerű (legfeljebb) évente megismételni, amennyiben a honlapon időközben bekövetkezett változások ezt indokoltá teszik.

Irodalomjegyzék

Breuer Pál, Dömölki Bálint, Havass Miklós (2025): A Data Archive about History of Informatics in Hungary, IEEE HISTELCON 2025 -Knowing the Past for Preparing the Future conference proceedings (megjelenés alatt)

Dömölki Bálint (2018): Structure of a historical data archive, Budapest (Kézirat) Hozzáférés: 2025.09.30. <https://real-ms.mtak.hu/27480/>

Havass Miklós (2024): Az iTF első 15 éve. NJSZT iTF., Budapest (Kézirat) Hozzáférés: 2025.09.30. <https://real-ms.mtak.hu/29425/>

- Holl, András (2013): *Az Akadémiai Könyvtár repozitóriuma, a REAL bővítése, gyarapítása, fejlesztése*. Tudományos és Műszaki Tájékoztatás, 60 (4). pp. 198-199. Hozzáférés: 2025.09.30. <https://real.mtak.hu/5830/>
- Magyar Informatikatörténeti Adattár. Megvalósítási rendszerterv. (2014). Budapest (kézirat) Hozzáférés: 2025.09.30. <https://ita.njszt.hu/objektum/magyar-informatikatortetneti-adattar/>