

DSD

Elosztott
Rendszerek
Osztály

Pataki Máté

MTA SZTAKI

- Bemutatkozás: MTA-SZTAKI Elosztott Rendszerek Osztály
- StreamOnTheFly - Új rádiózási formák
- A KOPI Plágiumkereső Portál
- Weblapok akadálymentesítése

MTA SZTAKI

A .hu velünk kezdődik

Elosztott
Rendszerek Osztály

DSD

- 1979. Az első hálózati kapcsolat a SZTAKI és az IIASA között
- 1982. A SZTAKI saját X.25 technológiát fejleszt ki, és ezzel később embargót tör
- 1983. Az első elektronikus levél külföldre
- 1986. Az első nemzetközi elektronikus levelezési kapcsolat UUCP/X.25 protokollal
- 1990. Az első bérelt vonalas kapcsolat a SZTAKI és a Linzi egyetem között
- 1991. Az első magyarországi gép az Interneten. A sztaki.hu az első magyar domain név és 192.84.225.1 az első IP cím.
- ...
- 1995. Ország első hálózati szótárszolgáltatása: SZTAKI Szótár
- ...
- 2002. A W3C megnyitja a Magyar Irodát a SZTAKI-ban
- 2003. W3C világkonferenciának ad otthont a SZTAKI

Elosztott Rendszerek Osztály

Bemutakozás - témakörök

Elosztott
Rendszerek Osztály

DSD

- Témakörök
 - Elosztott rendszerek, P2P
 - Webes rendszerek, Szemantikus Web szolgáltatások
 - Digitális könyvtárak, szövegfeldolgozás
 - Csoportmunka támogatás
- Európai uniós és hazai pályázatok, projektek
- W3C Magyar Iroda
 - Internetes szabványosítással foglalkozó szervezet
 - Egyetlen kelet-európai iroda
 - <http://www.w3c.hu>

Elosztott Rendszerek Osztály

Bemutakozás - projektek

Elosztott
Rendszerek Osztály

DSD

- NDA@SZTAKI
 - Nemzeti Digitális Adattár (<http://nda.sztaki.hu>)
- E-önkormányzati portál
 - Internetes ügyintézés
 - Bejövő levelek kiosztása (NLP)
- SZTAKI Szótár
 - Ingyenes online szótár (<http://szotar.sztaki.hu>)
- StreamOnTheFly
- KOPI Online Plágiumkereső és Információs Portál
- W3C-WAI, akadálymentesítés
- További információk
 - <http://dsd.sztaki.hu>

StreamOnTheFly - Új rádiózási formák

- Közösségi rádiók, internetes rádiók
- Műsorok egyszer elhangzanak, azután nyomuk vész
- Együttműködés, műsorcsere, globalizálódás
- Személyre szabott rádiózás
- StreamOnTheFly projekt: új módszerek keresése és kipróbálása



Fő gátló tényezők:

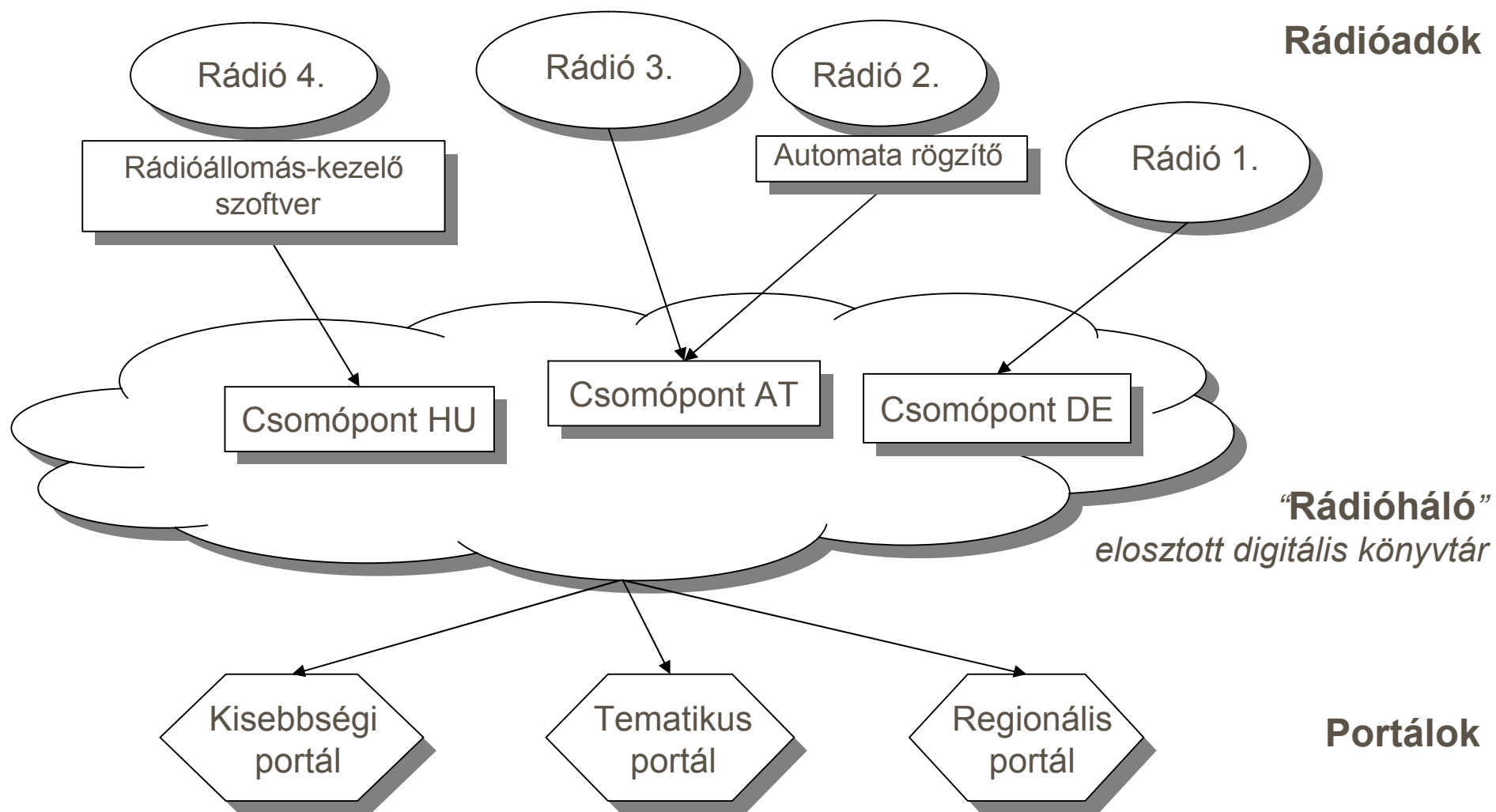
közös (meta)adatformátum hiánya

az együttműködés szoftveres támogatásának hiánya

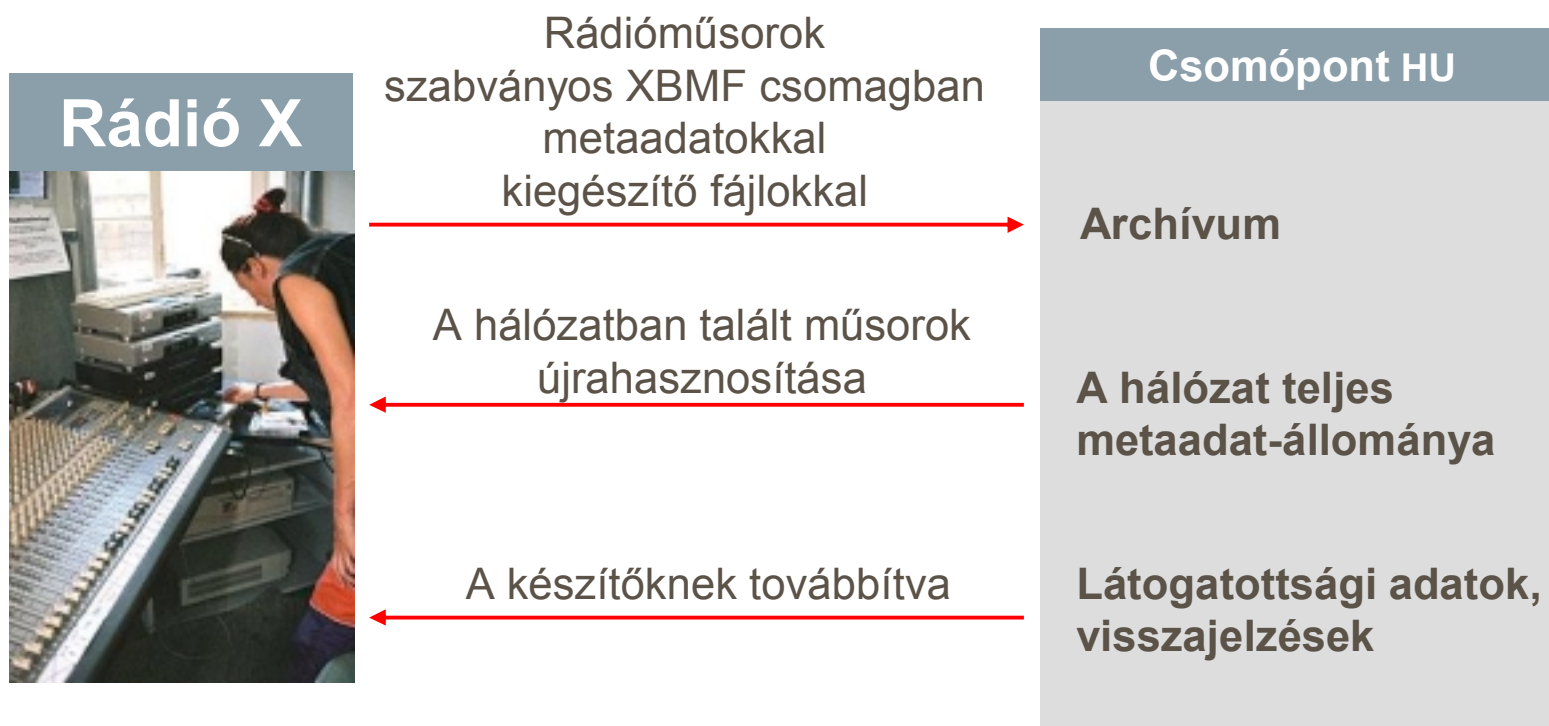
DSD

Elosztott
Rendszerek Osztály

A „rádióháló”



A rádióadó kapcsolata a csomóponttal



Programme Details

Programme Data

Programme Title:

*Date:

*Start Time:

*End Time: (if the programme goes over midnight, simply choose the morning end time)

*Special Needs:

*Programme Active:

*Options:  [Delete Programme](#)

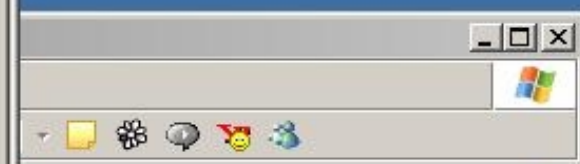
SOTF network meta data

Alternative Title:

*Keywords (comma separated):

*Programme Description:

Contributors (comma):

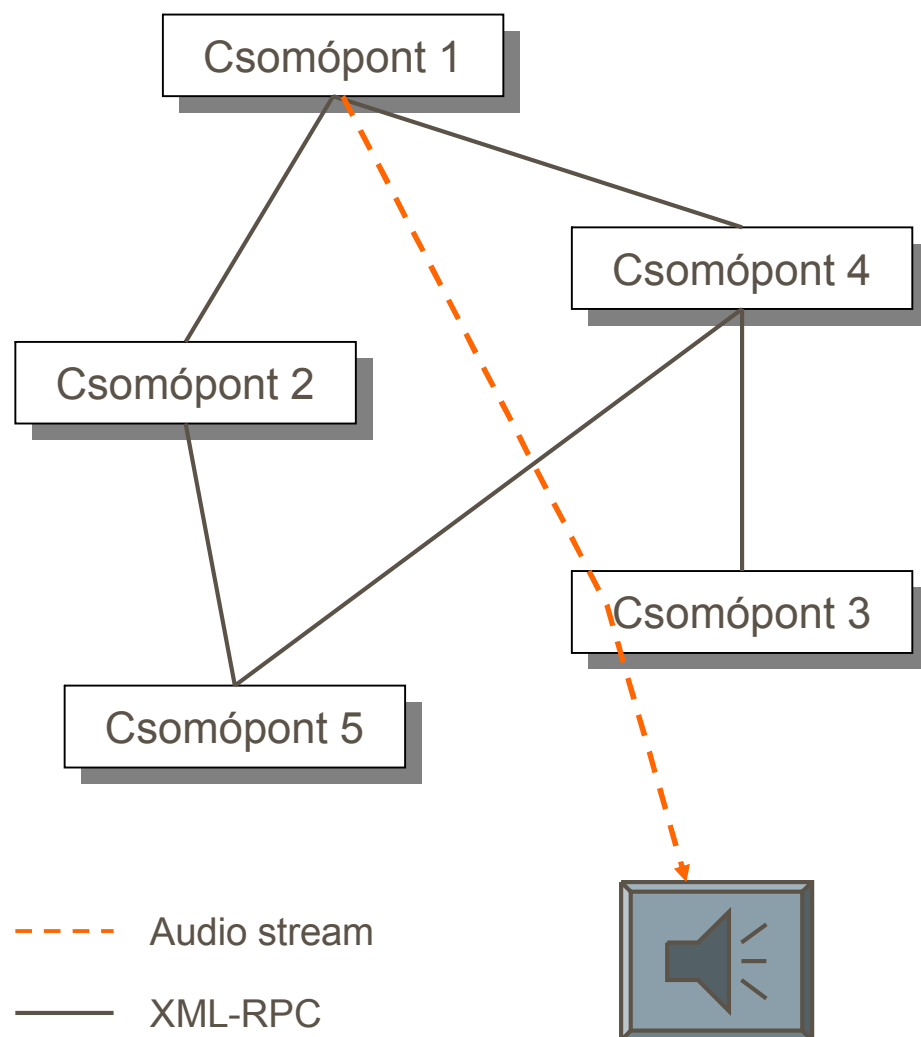


Calendar

<	April			2003		>
mo	di	mi	do	fr	sa	so
	01	02	03	04	05	06
07	08	09	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

8:00	
9:00	
10:00	 (Alex) Newz (10:30-11:00)
11:00	
12:00	
13:00	

A csomópontok hálózata



- Peer-to-peer hálózat
 - „szomszédok”
 - stream forwarding
- Egységes felhasználói felület
 - Böngészés
 - Témakörök
 - Összetett keresések
 - Kedvenclista, „önrádiózás”
 - Streaming
- Kényelmes szerkesztői felület



You are logged in as:
micsik
Preferences
Logout

Welcome to StreamOnTheFly! This is a network of audio archives containing radio programmes of community radios all over the World.

Wind, rain and forest



A big storm



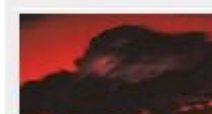
Nature



Loon in the nature



The peaceful ocean



It's electrifying

Any language

Search

ASD_Radio

Go

Topics with most programmes:

<u>Environment</u>	(18)
<u>Development</u>	(3)
<u>Information & Media</u>	(2)
<u>Health</u>	(2)
<u>Economy</u>	(1)

- untitled
- Why why why?
- WARNING: This document might be a little too complex
- Valami akarmi

3 radio stations
14 radio programmes
93 MB of audio
2 hours of audio
45 listens and
downloads
26 registered users

Az XBMF csomagformátumról

- Kell egy egységes módszer a műsorok tartalmának és metaadatainak együttes tárolására és mozgatására
- A létező ipari megoldások drágák és a célhoz képest túl bonyolultak (SMPTE, AAF, MPEG-7)
- A műsorhoz tartozó összes anyag egy csomagban:
 - hanganyag (többféle formátumban)
 - metaadatok
 - kiegészítő/háttér anyagok
- Metaadatok Dublin Core alapokon
 - (EBU Core, SOMA)
- Egyszerűen kezelhető, rugalmasan bővíthető, megkönnyíti a jelenlegi rendszerekről való áttérést (export-import)

Metaadatok

Cím, sorozatcím,
rövid leírás, kulcsszavak,
készítők, dátumok, jogok,
stb.

Kiegészítők

fényképek,
riport szövege,
stb.

Hangfájlok

mp3 128 kbps
mp3 24 kbps
ogg
stb.

XBMF = (Exchange Broadcast Binary and Metadata Format)

A portál kapcsolata a csomóponttal

Csomópont HU

Archívum

A hálózat teljes metaadatállománya

Látogatottsági adatok, visszajelzések

Az ajánlott műsorok meghallgathatók

Válogatás a hálózat műsoraiból:
egyesével (lista) vagy
keresőkifejezéssel (frissül)

Megjegyzések, értékelések



SOTF Portal - Microsoft Internet Explorer

Fájl Szerkesztés Nézet Kedvencek Eszközök Súlyó

Vissza Keresés Kedvencek Multimédia

SZUPERRADIO

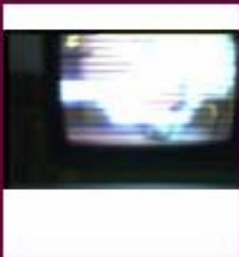
ONE GREAT STREAMONTHEFLY PORTAL

Listen to the
LOOPSTREAM of SZUPERRADIO
in 24 kbps MP3
or 64 kbps OGG



ooo **Laszlo marries a SOTF portal.**
Why we think this is the coolest way to present streaming audio on the web. Automated generation of Radio frontends - or Portals - without deep knowledge of web design. SOTF does it for you. [\[MORE\]](#)

LISTEN 24MP3 | 64OGG | 128MP3 | 3 COMMENTS | RATING ***



ooo **Television is evil.**
Why we think this is the coolest way to present streaming audio on the web. Automated generation of Radio frontends - or Portals - without deep knowledge of web design. SOTF does it for you. Why we think this is the coolest way to present streaming [...] [\[MORE\]](#)

LISTEN 24MP3 | 64OGG | 6 COMMENTS | RATING *



ooo **How our nodes work**
Why we think this is the coolest way to present streaming audio on the web. Automated generation of Radio frontends - or Portals - without deep knowledge of web design. SOTF does it for you. [\[MORE\]](#)

LISTEN 24MP3 | 128MP3 | 0 COMMENTS | RATING ****



ooo **Radio Aufdraht**
Why we think this is the coolest way to present streaming audio on the web. Automated generation of Radio frontends - or Portals - without deep knowledge of web design. SOTF does it for you. [\[MORE\]](#)

LISTEN 24MP3 | 64OGG | 128MP3 | 2 COMMENTS | RATING ***

Author name
Station name
Series name

Duration: 0:10:34
23-Oct-2002



DJ Superfun
NRJ 104,2
Das beste aus 3 Jahren

Duration: 1:00:00
22-Oct-2002



DJ Püringer
Radio FRO
Lustig samma

Duration: 1:00:00
22-Oct-2002



Christian Berger
orange 94.0
Radio Aufdraht

Duration: 1:00:00
22-Oct-2002



- A StreamOnTheFly rádióháló egy önszervező közösségtámogató szoftverkörnyezetet próbál kialakítani, amely remélhetően „önálló életre kel”, és új együttműködési formákat alakít ki a médiakészítés és sugárzás területén
 - Önszervező, decentralizált hálózat
 - Szabad szoftverek, nyílt protokollok használata
 - Rugalmasan kiegészíthető
 - Nem csak rádióra alkalmazható
- További információ:
 - <http://www.streamonthe-fly.org>
 - <http://dsd.sztaki.hu>



DSD

Elosztott
Rendszerek Osztály

Demó

<http://radio.sztaki.hu>

A KOPI projekt ismertetése

- KOPI Online Plágiumkereső és Információs Portál - internetes **hasonlóság** és **plágiumkereső** szolgáltatás
- Partner: Monash University, Melbourne
- 2002 ITEM pályázat (IHM-OM)
- MTA-SZTAKI Elosztott Rendszerek Osztály üzemelteti és fejleszti folyamatosan

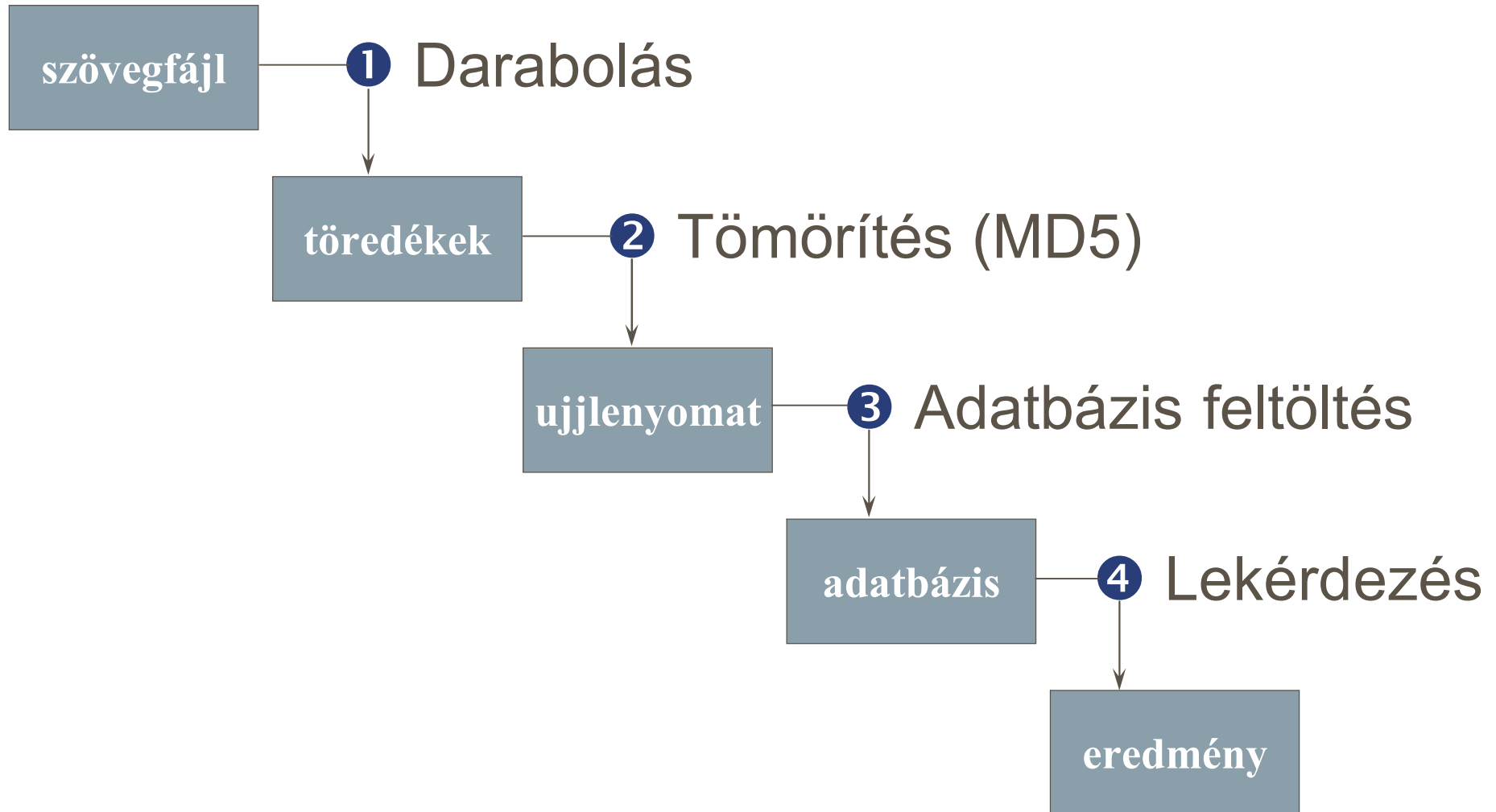
A KOPI projekt célja

- *Digitális könyvtárak védelme az illegális másolatoktól*
- *Tanárok, professzorok, konferenciaszervezők segítése a részben vagy egészében másolt dolgozatok illetve publikációk gyors megtalálásában és leleplezésében*
- *Diákok és cikkírók tájékoztatása a szabályos és helyes idézés formájáról és mértékéről (mértékének megállapítását a rendszer el is végzi)*
- *Diplomák és szakdolgozatok értékének növelése eredetiségük igazolásával*

Miért van rá szükség?

- Hasonló rendszerek
 - Teljes hasonlóságot néz
 - Írás stílusa alapján azonosítja a szerzőt
 - Tesztet generál a dolgozatból
 - Internetes keresőket használ
 - Ismeretlen belső működésű, fizetős rendszerek
- Ingyenes, nyelv-független algoritmus, magyar nyelvű felület, magyar tartalom

Másolatkereső rendszer működése



Szavas darabolás

- Eredeti

Ezen project célja, hogy a Monash University-vel együttműködve egy olyan rendszert hozzunk létre, amely hatékony a dokumentum-másolatok felderítésében.

- Szavas darabolással (n=5)

ezen project célja hogy a
monash university vel együttműködve egy
olyan rendszert hozzunk létre amely
hatékony a dokumentum másolatok felderítésében

Átlapolódó szavas darabolás

- Eredeti

Ezen project célja, hogy a Monash University-vel együttműködve egy olyan rendszert hozzunk létre, amely hatékony a dokumentum-másolatok felderítésében.

- Átlapolódó szavas darabolással (n=5)

ezen project célja hogy a

project célja hogy a monash

célja hogy a monash university

hogy a monash university vel

a monash university vel együttműködve

monash university vel együttműködve egy

university vel együttműködve egy olyan

...

Hash kódoláson alapuló tömörítés, MD5 algoritmus



- A bemenet bármilyen hosszú lehet
- Gyors
- Annak az esélye, hogy két különböző szövegnek ugyanaz a kódja, nagyon kicsi
- Irreverzibilis, veszteséges kódolás (biztonságos)


[Home](#)
[Laws and university
policies](#)
[Netiquette](#)
[Plagiarism detection demo](#)
[Forum](#)
[FAQ](#)
[Help](#)

Email address:

Password:

[Login](#)
[Login help](#)
[Create new account](#)

Plagiarism detection demo

You can test the system by inserting poems of the famous Hungarian poet Sándor Petöfi.
MEK

Megérte ezt az évet is,
Megérte a magyar haza;
A vész esen elborult,
De nem esett le csillaga.
Meg van vagdalva, vérzik a kezünk.
De azért még elbírja fegyverünk,
S amerre vág,
Ott hagyja fájó vérnyomát.

Buda várán újra német zászló!
Hova legyek lellem haragjától?
Arcunk piros, de nem zsarnok-vértől!
Hanem salát szívünk szégyenétől!

[Test](#)

Ujev napjan, 1849

 11%

Petőfi Sándor

Buda varan ujra nemet zaszlo!

 30%

Petőfi Sándor

Europa csendes, ujra csendes

 46%

Petőfi Sándor

Negy nap dorgott az agyu

 76%

Petőfi Sándor

Csataban

 88%

Petőfi Sándor

- Portál szolgáltatások
 - Témával kapcsolatos beszélgető-fórumok
 - Privát üzenetküldés, FAQ, nyelvváltás
 - W3C WAI irányelv
- Dokumentumok, ismertetőik
 - Plágium vagy idézet
 - Jogszabályok (szerzői jogok)
 - Egyetemi szabályzatok

Hasonlóságkereső szolgáltatások

- Feltöltött dokumentumok összehasonlítása egymással
- Hasonló dokumentumok keresése a rendszer adatbázisaiban
 - A felhasználó vagy mások által feltöltött dokumentumok
 - Internetről letöltött oldalak
 - Digitális könyvtárak (MEK)
 - Egyetemi diplomák
 - ...

- Elosztott rendszer
 - Minden egyetemen saját szerver, de
 - Egymás rendszerében is képesek keresni
 - Biztonságosan, MD5 számok átvitele
- Adatbázis feltöltése
 - Online és offline adatbázisok
 - Interneten található dokumentumok
- Forráskódok, programozási nyelvek felismerése

DSD

Elosztott
Rendszerek Osztály

Demó

<http://kopi.sztaki.hu>

- Nemzetközi „internetes szabványosítással” foglalkozó szervezet
- Olyan technológiák fejlesztésével foglalkozik, „melyek segítségével kiaknázzhatjuk a Web összes lehetőségét” (Tim Berners-Lee)
- W3C ajánlásai
 - HTML, XHTML, CSS, SVG, SMIL, MathML
 - XML, XSL
 - SOAP, WSDL, RDF, OWL
 - VoiceXML, SSML
 - ...
- Tagok irányítják a fejlődés irányát a munkacsoportokon keresztül

W3C Magyar Iroda Tevékenységei

- W3C fejlesztéseinek ismertetése itthon
 - Specifikációk, ajánlások, irányelvek, szoftverek
- Új tagok toborzása
 - Tagok alakíthatják a W3C és így a Web jövőjét
 - Fél-egy évvel hamarabb ismerik a következő szabványokat
 - Csökkent a tagdíj a magyar szervezetek számára
- Magyar nyelvű tartalom
 - Hírek
 - Ajánlások
 - Hírlevél
- Rendezvények szervezése

- MTA SZTAKI W3C tag
- <http://www.w3.org/WAI/>
- Web Content Accessibility Guidelines
 - 1.0 Recommendation (1999. máj. 5.)
 - 2.0: Working Draft (2004. nov. 19.)

- Web Content Accessibility Guidelines
 - Priority 1 (must) 
 - Priority 2 (should) 
 - Priority 3 (may) 
- Gépi és kézi ellenőrzés

Fogyatékosággal élők

- Fogyatékosággal élők
 - Vakok
 - Külön lap? [Budapest Portál](#)
 - Beszédes linkek
 - Gyengénlátók
 - Nagy kontraszt ([KOPI](#))
 - Betűméret állítása
 - Esetleg ábrák
 - Színvakok
 - Szín ne hordozzon fontos információt

Fogyatékosággal élők

- Fogyatékosággal élők
 - Hallássérültek
 - Zajos környezet (étterem, műhely)
 - Videó mellett felirat vagy leírás
 - Mozgássérültek
 - Csak egér vagy csak billentyűzet
 - Sok gépelés (cookie)
 - Értelmileg visszamaradottak
 - Vezetni a felhasználót (kikapcsolható)

Technológiailag megkülönböztetettek

- Technológiailag megkülönböztetettek
 - Eltérő képernyőméret
 - 19" LCD képernyők ([CNN](#) 220px)
 - Mobil, PDA (vízszintes görgetés)
 - Elavult böngésző
 - JavaScript, Flash, Java
 - Gyenge hardware
 - Modem (képek, Winword HTML)

Speciális célcsoportok

- Speciális célcsoportok
 - Gyerekek
 - Nem tud olvasni (Manó)
 - Sok választási lehetőség nem jó
 - Idősek
 - Fix betűk, mozgó ikonok, reklámok, popup
 - Alacsony (informatikai) képzettségűek
 - Súlygó (ne informatikus írja, context sensitive)
 - Idegen nyelvek és kultúrák
 - Cégeknel (idegen nyelvű vagy nyelven is)
 - Dátumok, fogalmazás (Amazon.com és .jp)

- Flash
 - Csak a nyitóoldal ([FKF](#))
 - Alternatív tartalom ([Harry Potter](#))
- JavaScript
 - Alternatív lehetőségek
 - Kliens oldali ellenőrzés
 - CSS és/vagy JavaScript: [DynDemo](#)
- Applet és ActiveX
 - Nincs mindig engedélyezve
 - Bizonyos esetekben szükséges (online játék)

Működés ellenőrzése

- Szöveges böngésző
- Opera
 - Szöveges böngésző
 - Stíluslap kikapcsolása
 - PDA
- Internet Explorer + PopupCop
 - Flash kikapcsolása
 - ActiveX kikapcsolása
 - JavaScript kikapcsolása

DSD

Elosztott
Rendszerek Osztály

Honlap

<http://www.w3c.hu>

Köszönöm a figyelmüket!

WEB: *<http://dsd.sztaki.hu>*

Email: *Mate.Pataki@sztaki.hu*