

ELOSZTOTT DIGITÁLIS KÖNYVTÁRI PROJEKT EURÓPÁBAN

Kovács László, laszlo.kovacs@sztaki.hu

Micsik András, micsik@sztaki.hu

MTA SZTAKI, Elosztott Rendszerek Osztály

Abstract

The member institutes of ERCIM (European Research Consortium for Informatics and Mathematics) decided to serve their technical reports through a unified and distributed digital library environment. This environment is based on the Dienst server of the American NCSTRL project. The server at this time is running in several computer research institutes in France, Sweden, Italy, Portugal, Greece and Hungary. It provides advanced search and display facilities, where the technical reports of selected institutions can be handled parallelly through a unified interface.

1. Egységesítési törekvések a könyvtári rendszerek világában

A közös európai digitális könyvtárak megvalósulásának napja egyre közelebb kerül. Ennek egyik jele az, hogy az európai számítástechnikai kutatóintézetek egy egységes, de elosztott rendszeren keresztül teszik elérhetővé kutatási beszámolóikat. Az ERCIM (European Research Consortium for Informatics and Mathematics) tagintézményei közül immár Franciaországban (INRIA), Svédországban (SICS), Olaszországban (CNR), Portugáliában (INESC), Görögországban (FORTH) és Magyarországon (SZTAKI) is működik a az elosztott könyvtári rendszer. Egy kis történelmi áttekintés után ennek a rendszernek a képességeit mutatjuk be.

Az elosztott digitális könyvtárak létrejöttének nélkülözhetetlen előfeltétele az egységes katalógus formátum. Immár jónéhány nemzetközi szabvány is van erre, melyek azonban sajnálatos módon nem teljesen kompatibilisek egymással. A fent említett európai kutatóintézetekben is több mint háromféle katalógusformátumot használnak. Külön csak az RFC1807 dokumentumban leírt katalógusformátumot említjük meg, mert ez tekinthető az Internet szabványának, és ezt használja az ERCIM elosztott digitális könyvtári projekt is.

A digitális könyvtárak két fő szolgáltatása a keresés és a kiválasztott dokumentumok megjelenítése vagy letöltése. Az amerikai WATERS (Wide Area Technical Report Service) szolgáltatás ezt úgy oldotta meg, hogy egy központi szerveren tárolta az összes hozzá tartozó dokumentum katalógusrekordját, és itt lehetett ezekben keresni. Maguk a dokumentumok mindig a származási intézmény szerverén helyezkedtek el, és innen voltak letölthetők. Hasonló megoldást jelent a World Wide Weben keresztül elérhető dokumentumok felderítése és leindexelése, annyi különbséggel, hogy itt a teljes szövegben kereshetünk.

Ha a szóba jövő technológiák szempontjából nézzük az elosztott digitális könyvtárak problémáját, a dokumentumok letöltésére két jól bevált eszköz is van: az FTP és a WWW. A keresést és a letöltést is megoldja a WAIS, vagy egy indexelt WWW szerver. Az igazi gond az egyes elosztott szerverek összefogása és egységes kezelése. Erre a célra hozták létre az amerikai NCSTRL projekt keretében a Dienst szervert, mely az ERCIM digitális könyvtári projektjének is kiindulópontját képezi. Az NCSTRL több mint 40 amerikai intézmény digitális könyvtárát egyesíti a Dienst segítségével.



**European Research Consortium
for Informatics and Mathematics**
ERCIM

MTA SZTAKI

ERCIM Technical Reports Library

Search the collection

- The [Fielded Search Form](#) allows you to perform a search on several fields of the bibliographic data, and/or to limit the search to specific institutions,
- Or simply enter [one or several words](#) into the box below to search author, title, and abstract:
- Or browse reports at any of the [participating institutions](#).

2. A Dienst szolgáltatás felépítése

A Dienst megközelítésében minden intézménynek saját magának kell ellátnia a keresés és a dokumentum megjelenítés feladatkörét is. Így a digitális könyvtár és a szükséges számítástechnikai infrastruktúra fenntartása is az intézmény feladata. A különálló szerverek összefogását egy úgynevezett metaszerver végzi, mely az összes szerver alapvető elérési adatait szolgáltatja.

Ez a felépítés az egyik oldalról megköveteli az állandó technikai felügyeletet (Dienst és WWW szerverek, valamint Internet kapcsolat állandó üzemeltetése), a könyvtárosi (kiadványok adatainak karbantartása) és szoftvermérnöki munkát (kiadványok formátumainak előállítás, konvertálás). A másik oldalról lehetővé teszi, hogy egy intézmény egy egységes keretrendszerben szolgáltatson dokumentumokat úgy, hogy közben teljes rendelkezési lehetősége van a dokumentumaival kapcsolatban, azaz elérhetővé teheti és visszavonhatja bármikor a dokumentumait, továbbá például korlátozhatja a hozzáférésüket, vagy egyedi szerzői jogokat tüntethet fel.

Szükséges még valamilyen hivatkozási séma kidolgozása a szolgáltatott dokumentumokra. A Dienst alatt minden dokumentumnak egyedi azonosítója van, mely alapján a dokumentumot egyszerűen meg lehet találni. Az azonosító is decentralizált elven épül fel. Az első része a szolgáltató intézmény (nevezük ezután kiadónak) azonosítója, mely központilag van bejegyezve, második része a kiadón belüli azonosító, melyet a kiadó rendel hozzá a dokumentumaihoz.

Bibliographic keywords: (AND keyword fields OR keyword fields)

Author:

Title:

Abstract:

Select one or more organizations from this list or search all organizations:

ERCIM

- CRIN (Centre de Recherche en Informatique de Nancy)
- ICS FORTH
- IMAG**
- INESC

NCSTRL

- Auburn University
- Boston University
- Caltech Computer Science
- Carnegie Mellon

Document Identifier:

A Dienst szerverek felhasználói felülete integrálva van a World Wide Web-bel, mely biztosítja a felület széleskörű elérhetőségét, platform-függetlenségét és egységes megjelenését. Ezen felül még az is igaz, hogy mindegyik szerver ugyanazokat a szolgáltatásokat nyújtja a felhasználónak a többi szerveren tárolt adatokra, mint a sajátjaira, így gyakorlatilag mindegy, hogy melyik szerveren veszem igénybe a felhasználói felületet. Hasonló módon egy dokumentumot, ha ismert annak azonosítója, bármely szerverhez kapcsolódva le lehet tölteni.

3. A Dienst szolgáltatásai a felhasználóknak

A keresési űrlapon a szerző, cím és kivonat mezők szavaira kereshetünk, az egyes keresési kulcsok között ÉS, VAGY kapcsolatot állíthatunk be. Kiválaszthatjuk, hogy a felsorolt kiadók közül melyeknél akarjuk elvégeztetni a keresést. Ezután a keresés párhuzamosan hajtódik végre a kiválasztott szervereken, a felhasználói felület csak az eredmények összefésülését végzi.

A kapott dokumentumlista alapján bármelyik dokumentumot megtekinthetjük. Elsőként a dokumentumhoz tartozó katalógusrekordot kapjuk meg. (A Dienst-et használják csupán katalógusrekordok szolgáltatására is.) Ez alatt a dokumentum elérhető formátumainak listáját látjuk, mivel egy dokumentumot többféle formátumban is lehet szolgáltatni. A formátumok többsége egyszerűen integrálható a Dienst-be. Az egy fájlból álló formátumok közül TeX, Word, RTF, Postscript, és bármely más használata megoldott.

Replication within Distributed Digital Document Libraries

László Kovács and András Micsik
TR95-1
Aug 28, 1995

Cooperative and interoperability aspects of distributed digital document libraries are discussed. The commitment of ERCIM institutions to the Dienst protocol raised the necessity of the development of general detailed architecture of Dienst. A new Dienst architecture augmented by replication service is suggested. Replication is described in details. An abstract Reference Model of distributed digital document libraries can be concluded as well.

How to view this document

- Display an **overview** of the document in one of the following formats.
 - [Overview of thumbnail pages](#)
- Display a **selected page** in one of the following formats (document has 8 pages).

inline gif image	▲	1	Display page
hi-resolution tiff image	▼		

- Display the **whole** document in one of the following formats.
 - [Microsoft Word](#) 835072 bytes.
 - [HTML](#) 1734 bytes.
 - [PostScript](#) 119630 bytes.
- [Print or download all or selected pages.](#)

Lehetséges szkennelt dokumentumokat képi formátumban szolgáltatni. Ez olyan esetekben ajánlott, amikor más szöveges formátumra nem oldható meg a konvertálás, például munkaerő hiányában, vagy különleges grafikai megjelenés esetén. Az oldalak képének egyes részleteit ki is tudja nagyítani a rendszer. Másik érdekes lehetőség a dokumentum logikai szerkezetének, tartalomjegyzékének megjelenítése, melyből könnyen tudunk a dokumentum kívánt részére ugrani. Az integrálható formátumok között van a HTML is.

Utolsóként az egyes kiadók dokumentumainak megjelenési év vagy szerző szerinti böngészésének lehetőségét említjük meg.

4. A Dienst szerver belső felépítése

A szerver belülről három alapszolgáltatásra, és opcionális tartozékokra oszlik. Az alapszolgáltatások közé tartozik a dokumentumtár, az indexelő, és a felhasználói felület. Ezek mindegyike adott eljárás-készlettel rendelkezik, melyek a Dienst saját protokollja szerint hívhatók meg a többi szerver által. Így amikor a felhasználó egy távoli szerver dokumentumának megjelenítését kéri, a Dienst szerver, melyhez a felhasználó kapcsolódott, a kérés teljesítéséhez a belső protokoll segítségével közvetlenül a távoli szerver dokumentumtárával kommunikál.

A szervert integrálni kell egy WWW szerverrel, azaz konfigurálni úgy, hogy a WWW szerver továbbítsa a felhasználó kéréseit a Dienst szerver felé.

A könyvtárosok munkáját megkönnyítő segédeszközök is vannak a Dienst-hez, amelyek új dokumentumok beillesztését, formátumkonverziókat, az adatbázis ellenőrzését és dokumentumok törlését teszik lehetővé.

Irodalomjegyzék

Jim Davis, Carl Lagoze: "Drop-in" publishing with the World Wide Web.

URL: <http://www.ncsa.uiuc.edu/SDG/IT94/Proceeding/Pub/davis/davis-lagoze.html>

The Dienst protocol and server, URL: <http://cs-tr.cs.cornell.edu/info/server.html>

Digital Libraries, Communications of the ACM, April 1995 Vol. 38, No. 4

SZTAKI Dienst szerver, URL: <http://www.sztaki.hu/dienst/>