

Az Ingenta és a Google együttműködése a teljes szövegű indexelésben

Az Ingenta, a világ egyik legnagyobb tudományos tartalomszolgáltatója bejelentette, hogy sikeresen üzembe helyezték a *Googlebot* keresőrobotot az Ingenta.com felületen elérhető több mint 6000 online közlemény teljes szövegének indexelésére. A Google az Ingenta.com felületen szabadon elérhető metaadatokat indexeli. Ez napi 20-30 ezer oldal feldolgozását jelenti, így már 5,4 millió Google-felhasználó számára biztosítja a cikkek címe, a kulcsszavak, a szerzők neve és a kivonatok szövege mellett a cikkek teljes szövegében történő keresést. Annak ellenére, hogy a program elsősorban a teljes szövegekhez készült, minden felhasználó az Ingenta.com ellenőrzése alá kerül. A Google-ben kereső olvasó korlátozás nélkül eljuthat az Ingenta.com weblapján a keresett cikk kivonatához, és választhat, hogy előfizet a teljes szövegű szolgáltatásra, vagy alkalmankénti díjat fizet.

A teljes szövegű keresését a CABI Publishing, a Professional Engineering Publishing, az FD Communication Inc. és az American Chemical Society is kísérleti üzemben teszteli. További kiadók is jelezték csatlakozási szándékukat.

/Advanced Technology Libraries, 33. köt. 6. sz. 2004. p. 4./

(J. É.)

Webarchiválással foglalkozó konzorcium alakult az Egyesült Királyságban

Az Egyesült Királyság Webarchiválási Konzorciuma (*UK Web Archiving Consortium = UKWAC*) az Egyesült Királyság válogatott honlapjainak archiválását tűzte ki célul, hogy az eljövendő nemzedékek számára is megmaradjanak az értékes kulturális

és tudományos források. A UKWAC tagjai: a British Library, a Közös Információs Rendszerek Bizottsága (Joint Information Systems Committee = JISC), a Nemzeti Levéltár (National Archives), a walesi Nemzeti Könyvtár (National Library of Wales), a skóti Nemzeti Könyvtár (National Library of Scotland) és a Wellcome Trust.

A program első szakasza két évig tart. Ez idő alatt hatezer weboldalt gyűjtését és archiválását tervezik. A konzorcium tagjai a honlapok tulajdonosainak engedélyével dolgoznak, és együttes erővel igyekeznek kidolgozni a válogatás szabályait. A munkamegosztást úgy szervezték, hogy mindenki a saját témáihoz kapcsolódó tartalmú, illetve című oldalakat gyűjti. Így például a British Library a nemzeti kultúrával és a történelmi eseményekkel foglalkozó honlapokat, a Wellcome az orvostudományi, a Nemzeti Levéltár a kormányzati területekhez kapcsolódó anyagokat. A skóti és a walesi Nemzeti Könyvtár a két nép kultúrájával és történelmével foglalkozó honlapokat kutatja fel, a JISC az Egyesült Királyság felsőoktatási intézményeinek információ- és kommunikációtechnológiai (ICT) projektjei, innovációi körében végez gyűjtést. Az infrastrukturális költségeket (hardver, szoftver, fejlesztés, fenntartás) egyenlő arányban osztják meg a konzorcium tagjai közt.

A gyűjtéshez a HTTrack nyílt forráskódú programot használják, az archiváláshoz pedig a PANDORA Digital Archival System (PANDAS) rendszert. Ez utóbbit az Ausztrál Nemzeti Könyvtár (National Library of Australia) fejlesztette ki, és alkalmazta (együttműködve más intézményekkel) az Ausztrál Nemzeti Webarchívum (*PANDORA Australian National Web Archive*) létrehozásához.

További információ: <http://www.webarchive.org.uk>; <http://pandora.nla.gov.au>

/Library Journal, 2004. július 2. <http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA430465/>

(Z. D.)