

**PRISTUP ARHIVIMA WEBA: RJEŠENJE PROJEKTA
PRISTUP NORDIJSKOM ARHIVU WEBA**

ACCESS TO WEB ARCHIVES: THE NORDIC WEB ARCHIVE
ACCESS PROJECT APPROACH

Svein Arne Brygfjeld
Nacionalna knjižnica, Mo i Rana, Norveška
svein.brygfjeld@nb.no

UDK / UDC 027.54(48):004

004.738.52

Preuzeti rad / Reprinted paper

Primljeno / Received: 12. 7. 2002.

Sažetak

Nacionalne ustanove u svijetu uviđaju važnost digitalnih dokumenata objavljenih na World Wide Webu kao vrijednog dijela sadašnje kulturne povijesti. Prikupljanje, zaštita i osiguranje pristupa toj velikoj zbirci informacija predstavlja iznimno važan izazov. Članak ukazuje na neke općenite aspekte tog područja, a ide i u dubinu opisujući pristup arhivima weba. Opisan je projekt Nordijski arhiv weba (The Nordic Web Archive). Taj je projekt ukazao na neka nova rješenja u omogućavanju korištenja arhiva weba i time, donoseći praktična iskustva, unaprijedio modele pristupa.

Ključne riječi: nacionalne knjižnice nordijskih zemalja, obvezni primjerak mrežne građe, pobiranje weba, arhiviranje weba, arhiv weba, model pristupa arhivu weba, Nordijski arhiv weba, projekt Pristup Nordijskom arhivu weba

Summary

National institutions in the international community have acknowledged the importance of digital documents published on the World Wide Web as an integral part of the current cultural history. Collecting, preserving and providing access to this vast collection of information place before us a very important challenge. This article points to some general field aspects, reaching further depths of the issue with its description of access to Web archives. The Nordic Web Archive project is described to some extent. The project has pointed to some alternative ways of enabling users to take benefit of the Web archives, and by offering some insights into their practical experiences has brought improvement to the current access models.

Keywords: national libraries of Nordic countries, legal deposit of Web resources, harvesting, archiving, Web archive, access to Web archive model, The Nordic Web Archive, The Nordic Web Archive Access Project

Pobiranje, arhiviranje i zakonodavstvo

Uvod

Zaštita dokumenata objavljenih na World Wide Webu (u daljnjem tekstu: web) trenutno predstavlja jednu od najvažnijih aktivnosti velikog broja knjižnica i arhiva u svijetu. Knjižnice i arhivi proučavaju različite aspekte tog izazova pa se može uočiti veći broj rješenja. Nacionalne knjižnice nordijskih zemalja (Danska,¹ Finska,² Island,³ Norveška⁴ i Švedska⁵) ulažu značajne napore u traženju svog načina zaštite weba. Budući da se radi o području koje omogućuje suradnju u istraživanju i razvoju, te su se knjižnice udružile kako bi zajednički pronašle rješenja. Svoj su projekt nazvale *Nordijski arhiv weba*, a unutar njega uspostavljen je i manji nordijski projekt koji ima za cilj pronalazjenje rješenja u području pristupa arhivima weba.

Projekt naziva *Pristup Nordijskom arhivu weba* (NWA - The Nordic Web Archive Access Project)⁶ pokrenut je krajem 2000. godine, a planira se da će biti završen sredinom 2002. Projekt financira Nordunet,⁷ a temeljni mu je cilj pronalazjenje sofisticiranog i korisniku prilagođenog rješenja za pristup velikim povijesnim arhivima. U projektu sudjeluju sve nacionalne knjižnice nordijskih zemalja.

Definicije

Pojedini izrazi koji se koriste u ovom članku upotrijebljeni su u njihovom uobičajenom značenju. Najvažnije izraze ipak valja definirati.

Arhiv weba (*Web archive*): povijesna, dobro organizirana zbirka dokumenata koji dolaze s weba

Pobiranje (*harvesting*): proces prikupljanja dokumenata s weba pri čemu se koristi softver-robot

Navigacija (*navigation*): kretanje među mrežnom građom korištenjem adresa koje određuju smještaj građe

Prebiranje (*browsing*): praćenje veza u mrežnim dokumentima

Metapodaci (*metadata*): podaci o podacima; izraz koji se često koristi kao sinonim za kataložni zapis

Zaštita weba u nordijskim zemljama

Svaka je od nordijskih nacionalnih knjižnica u određenoj mjeri radila na pobiranju ili prikupljanju dokumenata s weba. Tako su uspostavljeni više ili manje sofisticirani arhivi weba. Švedska je Kraljevska knjižnica izvršila nekoliko cjelovitih i sveobuhvatnih pobiranja švedske mrežne domene.⁸ Cilj je bio prikupiti svu građu

¹Vidjeti: danska Kraljevska knjižnica, <http://www.kb.dk>.

²Vidjeti: finska Sveučilišna i nacionalna knjižnica, Helsinki, <http://www.lib.helsinki.fi/english/index.htm>.

³Vidjeti: islandska Sveučilišna i nacionalna knjižnica, <http://www.bok.hi.is/english/third.htm>.

⁴Vidjeti: norveška Nacionalna knjižnica, <http://www.nb.no>.

⁵Vidjeti: švedska Kraljevska knjižnica <http://www.kb.se>.

⁶Vidjeti: Nordijski arhiv Web, <http://nwa.nb.no>.

⁷Vidjeti: Nordunet2, <http://www.nordunet2.org>.

koja se mogla pronaći na poslužiteljima, a koja sadrži podatke na švedskom jeziku ili o Švedskoj. Finska i Island vodile su se istom zamisli, dok Danska i Norveška zasad obavljaju ograničeno prikupljanje građe.

Promatrano s pravnog aspekta nije potpuno jasno nalaze li se poslovi pobiranja građe i uspostave arhiva weba unutar zakonskih ovlasti tih knjižnica. Dok je Norveška svojim zakonodavstvom osigurala takva prava nacionalnoj knjižnici, u Švedskoj je u vrijeme pisanja ovog članka prekinuta uspostava arhiva jer je utvrđeno da ne postoji zakonski temelj koji dopušta provođenje pobiranja i arhiviranja.

S kulturološkog aspekta, sve spomenute knjižnice smatraju veliku količinu podataka koja je trenutno dostupna na webu značajnim dijelom naše sadašnje kulturne baštine.

Zakon o obveznom primjerku

Kao što je spomenuto, pobiranje i arhiviranje weba mora imati uporište u zakonskim propisima. Norveška je 1989. godine donijela novi zakon o obveznom primjerku.⁹ On predstavlja značajnu preradbu prethodnog zakona i obuhvaća gotovo sve vrste podataka i nositelje podataka koje možemo zamisliti. Tim zakonom nisu obuhvaćeni samo tradicionalni dokumenti, poput papirnatih, već i glazba, pokretne i nepokretne slike te radijski i televizijski program. Korak naprijed predstavlja i uključivanje onog što nazivamo *elektroničkim dokumentima*. Danas se podrazumijeva da u obuhvat takvih dokumenta ulaze i dokumenti objavljeni na webu. U pojedinim zemljama, primjerice u Australiji, problematika se obveznog primjerka rješava kroz ugovore s nakladnicima. Ipak, uočava se međunarodni trend u smjeru kojim je na tom području krenula Norveška, a ustanovama je doista potrebno omogućiti rad na tako važnom zadatku.

Dugoročni repozitorij digitalnih dokumenta

U časopisu *Zeitschrift für Bibliothekswesen und Bibliographie*¹⁰ objavljen je detaljni prikaz rezultata projekta NEDLIB¹¹ kojemu je kao temelj repozitorija poslužio model OAIS.¹² Dugoročni digitalni repozitorij norveške Nacionalne knjižnice može poslužiti kao još jedan primjer takvog arhiva. U nastavku će taj arhiv biti ukratko opisan i poslužit će kao referenca za prikaz rješenja pristupa. Repozitorij je uspostavljen s nekoliko ciljeva, a kao najvažniji se ističu sljedeći:

- repozitorij mora biti velikog i proširivog kapaciteta,
- mora biti jedino mjesto pohrane svih digitalnih dokumenata u obliku pojedinačnih datoteka i baza podataka norveške Nacionalne knjižnice,

⁸Vidjeti: Kulturarw3, švedska Kraljevska knjižnica, <http://kulturarw.kb.se/html/kulturarw3.eng.html>.

⁹Act no. 32 of 9 June 1989 relating to The legal deposit of generally available documents, http://www.nb.no/html/legal_deposit.html.

¹⁰Vidjeti: *Zeitschrift für Bibliothekswesen und Bibliographie* 48, 3-4(2001).

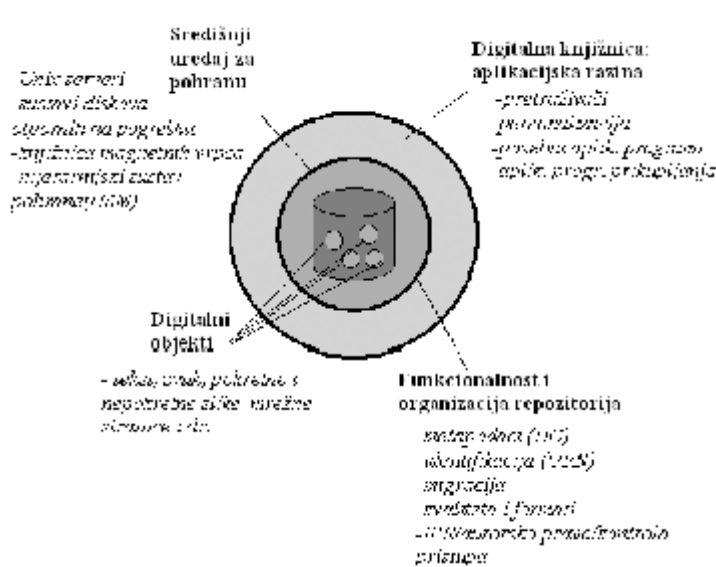
¹¹Vidjeti: Projekt NEDLIB, <http://www.kb.nl/coop/nedlib>.

¹²Vidjeti: OAIS Referentni model za otvoreni arhivski informacijski sustav, <http://ssdoo.gsfc.nasa.gov/nost/isoas/>.

- mora imati jedan mehanizam identifikacije koji æ globalno dodjeljivati jedinstvene i trajne identifikatore,
- mora podržavati migraciju dokumenata,
- mora biti neovisan o hardveru i omogućiti uporabu nove tehnologije u bilo kojem trenutku,
- mora koristiti otvoreni, standardizirani softver i operativni sustav.

Pojednostavljeni model takvog repozitorija prikazan je na Slici 1. Temelji se na zamisli da svi dokumenti budu smješteni u unutarnjem krugu dok se namjenski programi, kao što je pristup arhivu weba, nalaze u rubnom krugu. Prva inaèica repozitorija funkcionira veæ dvije godine kao centralizirana i opæa infrastruktura za upravljanje digitalnim dokumentima u norveškoj Nacionalnoj knjižnici.

Stoga postupak èuvanja digitalne zbirke postaje djelotvornijim, a usto je i precizniji jer se ne upravlja materijalnim entitetima pa je lakše pratiti i automatizirati svaki pojedini korak.



Slika 1

Pobiranje weba

Mnogi arhivi i knjižnice u svijetu ulažu znaèajne napore u pobiranju sadržaja weba. Ne treba zaboraviti da slièan posao veæ godinama obavljaju pretraživaèi weba, tako da nema sumnje u ostvarivost takvog zadatka. Kao najèešæe korišteni sustavi za pobiranje u nordijskim nacionalnim knjižnicama istièu se: *Combine*

harvester,¹³ *NEDLIB harvester*¹⁴ i *HTTrack*.¹⁵ Ipak, otkrivanje načina za sustavno prikupljanje dokumenata s weba spada u područje otvoreno istraživanju i razvoju. Primjerice, netko može poželjeti prikupiti što više inačica pojedinih online službi na webu, na primjer novina koje izlaze na internetu, i na taj način ukazati na brzu promjenu podataka. Čini se da rješenje takvog problema ovisi o tješnjoj suradnji nakladnika i ustanova koje vrše pobiranje jer korištenje sadašnje tehnologije ne omogućuje takvu vrstu pobiranja. Različiti aspekti pobiranja i arhiviranja detaljnije su opisani u časopisu RLG Digi News.¹⁶

Pristup mrežnom arhivu – rješenje projekta Nordijski arhiv weba

Temeljni ciljevi

Kao što je prethodno opisano, projekt Nordijski arhiv weba dio je šire suradnje na zaštiti weba u nordijskim nacionalnim knjižnicama. Usmjeravajući pozornost rješavanju problema pristupa, projekt će omogućiti:

- opće i dobro definirano sučelje arhiva,
- pristup utemeljen na pretraživanju, navigaciji i prebiranju,
- prostornu i vremensku podršku navigaciji,
- prenosivost među različitim rješenjima arhiva,
- povezivanje s drugim dijelovima digitalne knjižnice.

Druga područja od posebnog interesa uključuju distribuirane mehanizme za osiguranje pristupa većem broju arhiva kao i tehnologiju koja će podržati zahtjeve nacionalnih jezika te višejezični pristup.

Arhitektura

Odabrana arhitektura teži održavanju prenosivosti i promjenljivosti te daljnjem razvoju pojedinih dijelova implementacije. Kako je već spomenuto, sustav pristupa bit će neovisan o arhivu. Ta se odluka svakako odražava na arhitekturu sustava. Arhitektura i njezini dijelovi prikazani su na Slici 2. Uz nju su navedena objašnjenja, a procesi u samoj arhitekturi uključuju sljedeće:

- prihvatanje dokumenata: sustav pristupa primat će dokumente u formatu XML¹⁷ iz arhiva na sadržajno označavanje; ti dokumenti sadrže metapodatke u formatu Dublin Core, predviđajući buduću podršku za OAI (Open Archive Initiative);¹⁸ također, dostupni su i neki arhivski metapodaci, primjerice, vrijeme pobiranja;
- sadržajno označavanje: organizacija dokumenata za pretraživanje;
- pretraživač: element koji podržava pretraživanje indeksiranih dokumenata i metapodataka;

¹³Vidjeti: The combine harvester/indexes, <http://www.lub.lu.se/combine>.

¹⁴Vidjeti: The NEDLIB harvester, <http://www.csc.fi/sovellus/nedlib>.

¹⁵Vidjeti: The HTTrack Website Copies, <http://www.httrack.com/page0.php>.

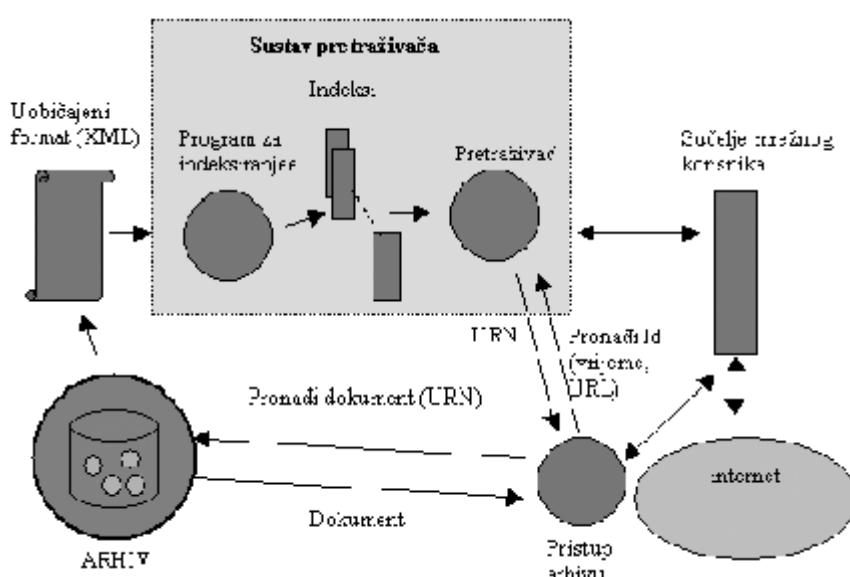
¹⁶Vidjeti: RLG DigiNews 5, 2(2001), <http://www.rlg.org/preserv/diginews/diginews5-2.htm>.

¹⁷Vidjeti: Extensible Markup Language, <http://www.w3.org>.

¹⁸Vidjeti: Open Archives Initiative, <http://www.openarchives.org>.

- mrežno suèelje: element koji omogućuje korisnièko suèelje zasnovano na korištenju weba;
- pristup arhivu: element za podršku dostave dokumenata iz određenog arhiva weba.

Rezultat je takve arhitekture da se arhiv i sustav pristupa u naèelu mogu promatrati kao dva odvojena sustava. Ona također omogućuje daljnja proširenja, primjerice primjenu Z39.50 ili suèelja temeljenih na OAI-u.



Slika 2

Naèini pristupa

Promatrano s korisnièkog aspekta, postoji nekoliko naèina pristupa arhivu weba. Arhivi u nordijskim nacionalnim knjižnicama trebali bi predstavljati sliku onog što se na webu nalazilo u određenom trenutku. To znaèi da je razlika između arhiva i samog weba u èinjenici da arhiv može sadržavati više inaèica određenog dokumenta, a može uključivati i dokumente koji više ne postoje na webu.

Temeljna naèela

Prije svega, sadržaj weba u arhivu izgledat æe isto kako je izgledao na mreži. To se naèelo èini prilièno jasnim i, kako je reèeno, navedeni arhivi pružaju zrcalnu sliku onog što se može pronaæi na webu.

Drugo, pristup arhivu temeljit æ se na obiènom pregledniku weba tj. zahtjevi hardvera i softvera bit æ potpuno isti kao da se radi o pristupu dokumentima na mreži. Razlog tome jest želja da korisnik pristupa arhivu weba na internetu bez potrebe za bilo kakvim dodatnim softverom.

Još je važnije da se korisniku omoguæe jasni naèini pristupa. Potrebno je težiti uporabi korisniku poznatih naèina pristupa poput prebiranja, navigacije i pretraživanja. Sustav pristupa Nordijskog arhiva weba zasnovan je upravo na tim temeljnim funkcijama.

Pretraživanje

Za razliku od projekta Wayback Machine o kojem se u posljednje vrijeme mnogo govori, sustav pristupa Nordijskog arhiva weba omoguæit æ korisnicima pretraživanje.¹⁹ Prije svega, to znaèi da æ biti omoguæeno pretraživanje sadržaja dokumenata u arhivu. Usto, korisnik æ moæi kao temelj za pretraživanje koristiti i različite vrste metapodataka.²⁰ Korisniku æ biti omoguæeno i da, navoðenjem poèetnog i završnog datuma, ogranièi upit na odreðeno razdoblje.

Funkcionalnost pretraživanja zavisi od sustava odabranog pretraživaèa, a o kvaliteti pretraživaèa zavisi moguænost lijevog i desnog skraæivanja i pretraživanja prema *frazi*. Pretraživaè na kojem se temelji projekt omoguæuje desno skraæivanje rijeèi i pretraživanje prema *frazi*.

Kretanje kroz vrijeme

Jedna od uoeljivih dodatnih usluga arhiva weba jest moguænost pristupa različitim inaèicama odreðenog dokumenta. To je posebno znaèajno za podatke koje na webu nude online sluþbe. Primjer su za takvu sluþbu novine objavljene na mreži koje identificira i odreðuje im smještaj jedinstvena adresa graðe (Uniform Resource Locator - URL). Pristupom online sluþbi u različita vremena korisnik moþe zateæi potpuno nove sadržaje.

Sadašnja primjena korisnièkog suèelja koje podrþava kretanje kroz vrijeme temelji se na uporabi vremenske crte. Vremenska crta kao pomagalo za kretanje prikazana je na Slici 3 (istaknuta je elipsom). Na vremenskoj je crti oznaèeno koje su inaèice dokumenta dostupne pa se korisnik moþe lako kretati od jedne do druge. Jednostavnim pritiskom miša na oznake na strelici korisnik se pomièe na prethodnu ili sljedeæu inaèicu dokumenta. Kao što je vidljivo na slici, sam je dokument prikazan u donjem dijelu preglednika dok je gornji dio namijenjen navigaciji.

¹⁹Vidjeti: The Wayback Machine, <http://www.archive.org>.

²⁰Opširnije o metapodacima vidjeti u nastavku teksta.



Slika 3

Kretanje kroz prostor

Odredivši jednu točku u vremenu, korisnik se može kretati od jednog *mjesta* na drugo kako bi pogledao sadržaj kojega drugoga mrežnog dokumenta objavljena u isto vrijeme. Kretanje od jednog do drugog dokumenta odvija se tako da korisnik upiše izvorni URL traženog dokumenta. U elipsi na Slici 4 prikazano je mjesto na koje korisnik može upisati URL traženog dokumenta. Ako se traženi dokument nalazi u arhivu, sustav ga povlači i dostavlja korisniku.



Slika 4

Prebiranje

Kao i u izvornim dokumentima na internetu, i u arhivu æ biti moguæ pratiti veze meðu pohranjenim dokumentima. Na taj æ naèin dokumenti na koje upuæuje veza biti dostavljeni iz arhiva. Prebiranje podrþava jedan od vjerojatno najèešæih naèina korištenja weba: samo odaberi i *klikni* i nešto æ se dogoditi!

Metapodaci

– Definicije

U kontekstu arhiva weba i mreþnih dokumenata izrazom *metapodaci* oznaævavaju se razlièite vrste opisa. Prije svega, nakladnik dokumenta moþe izraditi podatke koji opisuju sadržaj dokumenta i koje druge formalne pojedinosti. Takve æu metapodatke u nastavku zvati korisnièkim metapodacima. Drugo, protokol za prijenos mreþnih dokumenata (*http – hypertext transfer protocol*) takoðer omoguæuje pristup odreðenim metapodacima. Za njih æu koristiti naziv metapodaci protokola. Treæe, neki su metapodaci izraðeni radi olakšavanja arhivske funkcije dokumenata. Te æu metapodatke zvati arhivskim metapodacima.

– Dostupnost korisnièkih metapodataka

Èini se da veæina nakladnika ne izraðuje relevantne metapodatke za svoje dokumente. Osim toga, ima razloga da takvim metapodacima ne trebamo previše vjerovati. Primjerice, dogaða se da oni koji izraðuju razlièite metapodatke za svoje dokumente, nastoje njima manipulirati kako bi postigli veæu pretraþivost ili više rangiranje dokumenata u opæim pretraþivaèima. Ustanove koje vrše pobiranje mreþnih dokumenata mogu same izraðivati korisnièke metapodatke tj. kataloþne zapise. Zbog velike kolièine dokumenata objavljenih na mreþi ne moþe se oèekivati da æ dokumenti prikupljeni pobiranjem sadržavati relevantnu razinu opisa. Iako HTML nije jezik koji je namijenjen opisu strukture dokumenta, on sadrþi neke oznake polja koje se mogu iskoristiti za automatsku izradu metapodataka. Primjerice, oznaka polja *<NASLOV>* upuæuje na èinjenicu da je tekst koji slijedi iza njega naslov dokumenta. Korištenje moguænosti ogranièenja pretraþivanja na odreðene oznake polja korisnicima moþe omoguæiti dodatnu kvalifetu pretraþivanja.

Dostava dokumenata

Buduæi da je sustav pristupa neovisan u odnosu na arhiv, ne moþe se ni osloniti na arhiv za dostavu dokumenata. Referentni model identificira modul èija je funkcija da èini suèeelje izmeðu sustava pristupa i arhiva. Taj je model u obliku cgi-programa koji se aktivira u trenutku kad mreþni posluþitelj pošalje neki zahtjev.

Identifikacija dokumenata

Dostava određenog dokumenta podrazumijeva i proces identifikacije i određivanje smještaja tog dokumenta. Budući da se dokumenti dostavljaju iz arhiva, njihovi identifikatori mogu biti bilo što. Ipak, kako bi se postigla potpuna fleksibilnost i što manja zavisnost od arhiva, sustav pristupa pretpostavlja da su URL određenog dokumenta kao i URL-ovi njegovih unutarnjih poveznica valjani samo u mjesnom sustavu pristupa. Kako bi se omogućila usklađenost između sustava pristupa i naziva u arhivu, sustav očekuje da arhiv osigura usklađenost ta dva sustava naziva. Kad se uspostavi takva usklađenost, moguće je bolje iskoristiti prednost brzine pretraživača kako bi se omogućio brz pristup arhivu. Za 100.000 dokumenata formata HTML u arhivu weba Nacionalne knjižnice Norveške uspostavljeno je u prosjeku 15 vrsta unutarnjih odnosa (veze na slike, okvire, multimedij i drugo) koji moraju biti određeni kako bi se dokument mogao dostaviti korisniku.

Korištenje pretraživača

Za projekt Nordijski arhiv weba odabran je komercijalni pretraživač koji omogućuje svu funkcionalnost pretraživanja u sustavu. Više je razloga takvog odabira:

- posao implementacije ograničen je na preostale dijelove sustava čime projekt postaje realističniji, a složenost implementacije izrazito se smanjuje,
- takav pretraživač pokazuje izvrsnu izvedbu pretraživanja: dobra izvedba nudi mogućnost ispitivanja novih korisničkih slučajeva,
- poslove oko najkompleksnijih dijelova sustava obavlja netko drugi: implementacija sofisticiranih funkcionalnosti pretraživanja koje su visoke kvalitete izvedbe izuzetno je težak zadatak,
- projekt se zato može usmjeriti važnijim područjima: postizanju dobre fleksibilnosti slučajeva arhiva weba i ispitivanju novih korisničkih slučajeva što je od iznimne važnosti za knjižnice i arhive.

Za ovaj je projekt odabran softver Fast Search&Transfer²¹ jer je najbolji opći pretraživač. Sadrži pomagalo za sadržajno označavanje i pretraživač, a izuzetno je učinkovit u izvođenju operacija i na sustavu manjih mogućnosti zasnovanom na Linuxu. Odabir tog softvera pokazao se dobrim i za distribuiranu arhitekturu kojom se može postići veća izvedba te za distribuirana pretraživanja.

Open Source

Već je rečeno da je jedno od temeljnih načela uspostave sustava bilo postizanje njegove neovisnosti od arhiva što omogućuje da se implementacija može koristiti u različitim okruženjima. Drugi ključan problem prenosivosti vezan je uz zavisnost od tuđeg softvera. Kako je trenutno rješenje uglavnom zasnovano na standardnoj Linux inačici, problem prenosivosti bit će riješen. Izazov predstavlja i zavisnost od pretraživača koji je bitan za funkcionalnost sustava. U posljednju je inačicu uloženo

²¹Vidjeti: Fast Search&Transfer, <http://www.fast.no>.

puno truda kako bi se dobro definiralo suèelje pretraživaèa. To znaèi da æe prijelaz na drugi pretraživaè biti moguć uz ogranièena ulaganja.

Kao doprinos razvoju funkcionalnih rješenja u ovom području, konaèna æe implementacija projekta biti dostupna na internetu kao *Open Source*.²² Nadamo se da æe se na taj naèin i drugi prikljuèiti radu na poboljšanju rješenja koje je izradila *virtualna projektna grupa* sastavljena od istraživaèa svih ustanova ukljuèenih u projekt.

Iskustvo i izazovi

Dosadašnja iskustva projekta Pristup Nordijskom arhivu weba ukazuju da je uspostava zadovoljavajuæeg sustava za pristup arhivima weba prilièno ostvariva. Projekt æe osigurati poèetni softverski temelj na kojem se mogu graditi slièni sustavi, ali zasigurno æe se pojaviti i neka druga rješenja. Znaèajni izazovi postoje u još jednom području povezanom s pristupom arhivu weba. Promatrano s pravnog aspekta, još uvijek nije potpuno jasno je li dopušteno arhive weba uèiniti dostupnima javnosti na internetu. Prava intelektualnog vlasništva mogu takav pristup svesti na minimum, a veæina pristupa morala bi ovisiti o ugovorima. Projekt se uopæe nije bavio tom problematikom. Osim toga, èini se da glavni izazov predstavlja pobiranje weba i izgradnja velikih arhiva weba. Meðu tim izazovima pozornost osobito treba usmjeriti na:

- duboki web (*Deep Web*): podaci dostupni na mreži kojima se iskljuèivo može pristupiti putem korisnièkog upita,
- dinamiène i korisnièki uvjetovane sluþbe,
- neprekidnost: kako osigurati cjelovitu sliku svih izmjena dokumenata koji se izuzetno brzo mijenjaju,
- automatska izrada metapodataka: digitalne dokumente uglavnom neæemo moæi opisati na naèin kako to èinimo s tradicionalnim dokumentima.

*(Prevela Sofija Klarin,
struèna redakcija prijevoda Mirna Willer.)*

Napomena

Èlanak je izvorno objavljen u èasopisu *Zeitschrift für Bibliothekswesen und Bibliographie* 49, 4(2002), 227-231. Prijevod èlanka objavljujemo uz dopuštenje autora i Uredništva spomenutog èasopisa.

²²Vidjeti: Open Source Initiative, <http://www.opensource.org>.