

RDF: MOGUĆNOSTI PRIMJENE

RDF: POSSIBILITIES OF APPLICATION

Koraljka Golub

Katedra za bibliotekarstvo, Odsjek za informacijske znanosti
Filozofski fakultet u Zagrebu

kgolub@ffzg.hr

Rober Ravnić

Nacionalna i sveučilišna knjižnica, Zagreb

rravnic@nsk.hr

Sažetak

Članak obrađuje RDF - Resource Description Format na kojem se počelo raditi prije četiri godine zbog potrebe za modelom koji bi podržavao uporabu različitih formata metapodataka. RDF osigurava interoperabilnost među aplikacijama koje razmjenjuju računalno čitljive informacije na World Wide Webu. Razvija ga Konzorcij World Wide Weba (World Wide Web Consortium), a na snazi je prva inačica preporuke iz 1999. godine. RDF omogućava svakoj pojedinoj zajednici da sama definira svoj rječnik u obliku tzv. RDF sheme, a XML sintaksa osigurava i jednoznačnost rječnika. Mogućnosti su primjene RDF/XML-a raznolike, od opisa i pronalaženja izvora, rangiranja izvora, upravljanja osobnim podacima do digitalnih potpisa. Pokrenuti su mnogi projekti koji razvijaju alate za konverziju različitih formata metapodataka i izradu sučelja za pretraživanje vezanih uz RDF/XML, a određen broj web-stranica primjenjuje DC (Dublin Core) RDF/XML za opis izvora kojima omogućava pristup.

Cljučne riječi: RDF, XML, primjena, DC, RDF shema, XML namespace

Summary

The article deals with Resource Description Format (RDF), the development of which started four years ago in order to provide a model which would support a variety of metadata formats. RDF ensures interoperability between applications exchanging machine-readable information on the World Wide Web. RDF is being developed by the World Wide Web Consortium, and the current version is the 1999 recommendation. It enables each community to define its own vocabulary in the form of the so-called RDF scheme, while XML syntax ensures its uniformity. Possible applications of RDF/XML system range from resource description and discovery, content rating and privacy management to digital signatures. A number of projects have been started, developing tools for conversion of different metadata formats and designing search interfaces related to RDF/XML, whereas a number of web sites already apply Dublin Core RDF/XML for resource description to which they provide access.

Keywords: RDF, XML, application, Dublin Core, RDF scheme, XML namespace

1. Uvod

RDF (Resource Description Framework) zamišljen je kao model metapodataka koji podržava uporabu različitih formata metapodataka, a temelji se na sintaksi XML-a (eXtensible Markup Language). RDF omogućava interoperabilnost među aplikacijama koje razmjenjuju strojno čitljive informacije na webu iz svih područja ljudskog djelovanja.

Razvija ga Konzorcij World Wide Weba (World Wide Web Consortium: W3C) od 1997. godine. Temelji se na iskustvima nekoliko zajednica, uključujući sam Konzorcij (HTML metapodaci i PICS - Platform for Internet Content Selection), knjižničnu zajednicu, zajednicu koja se bavi strukturom dokumenata (SGML i XML) te zajednicu koja se bavi prikazom znanja. Prva i dosad jedina inačica službene preporuke objavljena je početkom 1999. godine pod nazivom Resource Description Framework (RDF) model and syntax specification.¹ Uloga je Konzorcija da promiče uporabu ovoga dokumenta, s ciljem pospješivanja funkcionalnosti i interoperabilnosti weba.

U protekle četiri godine pokrenuti su mnogobrojni projekti. Jedan je od najznačajnijih OCLC-ov projekt kooperativne katalogizacije mrežnih izvora - CORC (Cooperative Online Resource Catalog).² OCLC, član Konzorcija, drži da će RDF/XML postati osnova za razmjenu metapodataka na webu te ga intenzivno koristi. Zapisi se mogu izraditi u MARC-u ili DC-u, a prikazati u MARC-u, DC HTML-u i DC RDF/XML-u. Određen broj projekata bavi se razvojem tehničke infrastrukture za razmjenu metapodataka putem RDF/XML-a, što se odnosi na specifikaciju RDF shema, pomagala za prevođenje formata i sučelja za pretraživanje. Projekt UK Online,³ mrežno mjesto za britanske službene informacije na webu, opisuju izvore u DC RDF/XML-u.

2. Primjena RDF-a

RDF može biti osobito koristan pri pronalaženju informacijskih izvora jer poboljšava mogućnosti tražilica (*search engines*). Nadalje, u RDF/XML-u opisuju se sadržaji web-stranica, kao i odnosi među pojedinim sadržajima unutar njih, kako jedna stranica "proizlazi" iz prethodne i slično. RDF koristi i inteligentnim agentima za olakšavanje razmjene znanja, a omogućuje i rangiranje sadržaja.⁴ Primjenu nalazi i u opisivanju prava na privatnost, primjerice, kako stranice upravljaju osobnim podacima, hoće li otkriti te podatke drugima i što će korisnik dobiti zauzvrat. RDF-om se mogu opisati hardverske i softverske specifičnosti korištenja izvora za osobe s posebnim potrebama. Naglašava se i vjerovanje da će RDF imati posebnu važnost u uspostavljanju "weba povjerenja" za potrebe elektroničke trgovine i dr. zbog velike mogućnosti primjene u digitalnim potpisima.

¹Resource Description Framework (RDF) model and syntax specification. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222> (2001-10-20).

²OCLC CORC : web/based metadata creation for describing electronic resources. <http://purl.oclc.org/corc/> (2001-11-15)

³UK Government Information Services. <http://www.ukonline.gov.uk/> (2001-11-15).

⁴Isto.

3. RDF i XML

RDF koristi sintaksu XML-a za razmjenu i prijenos metapodataka. XML je metoda stavljanja strukturiranih podataka u tekstualnu datoteku, što omogućava prikaz podataka neovisno o programu koji ih je proizveo. Predstavlja skup pravila i smjernica za dizajn tekstualnih formata za strukturirane podatke, pomoću kojih se stvaraju datoteke koje računalo može jednostavno generirati i čitati budući da su nedvosmislene, neovisne o platformi te podržavaju internacionalizaciju i lokalizaciju. Budući da se i XML bavi strukturom podataka u dokumentima, postavlja se pitanje zašto uopće koristiti RDF. XML ne propisuje i ne zahtijeva određeni način opisivanja podataka tj. pojmovi se u XML-u mogu napisati na više načina od kojih će svi biti ispravni.

RDF je razvijen zbog potrebe za prepoznavanjem traženih podataka u strukturama formata. Ako ne može pretražiti format (npr., PDF datoteku), pretražit će priljepke na toj datoteci (npr., tipa vCard).⁵

U okviru RDF-a neovisne zajednice mogu razviti rječnike koji odgovaraju njihovim specifičnim potrebama i dijeliti ih s drugim zajednicama. Da bi se rječnici mogli zajednički koristiti, pojmove je potrebno detaljno definirati i opisati. Opisi takvih skupova pojmova nazivaju se RDF shemama.⁶ U RDF/XML-u shema se navodi u oznaci za imenski prostor (`xmlns = XML namespace`) zajedno s web-adresom specifikacije sheme (vidjeti primjer u sljedećem poglavlju). Različite sheme mogu imati iste oznake, primjerice, DC i različite vlasničke sheme nakladnika mogu imati istu oznaku "title", no koja se shema primjenjuje definirana je u imenskom prostoru.

4. Model RDF-a

Osnovni model RDF-a sastoji se od tri vrste objekata: izvora, svojstva i naredbe. Budući da se poistovjećuje s "prostom rečenicom",⁷ izvor je tj. rečenični subjekt sve što ima URI (Universal Resource Identifier), od ukupnih web-stranica, pojedinačne web-stranice, do predmeta koji se ne nalaze na webu, npr., audio-kaseta. Svojstvo (objekt) predstavlja određeni aspekt, atribut, karakteristiku nekog izvora ili odnos prema drugim izvorima. Svojstvo služi za određivanje značenja, dopuštenih vrijednosti i odnosa. Naredbu (predikat) čini izvor s pridruženim svojstvom i vrijednošću (*value*), koja se upisuje uz to svojstvo za taj izvor. U primjeru koji slijedi, web-stranica pod naslovom *Using Dublin Core* "http://

⁵Jane Hunter i Carl Lagoze u članku *Combining RDF and XML schemas to enhance interoperability between metadata application profiles* predložili su mehanizme za interoperabilnost metapodataka iz različitih zajednica na temelju RDF i XML shema. Projekti za interoperabilnost metapodataka različitih zajednica temelje se na RDF shemama ili XML shemama. Vidjeti: <http://archive.dstc.edu.au/RDU/staff/jane-hunter/www10/paper.html>. Primjerice, projekt SCHEMAS koristi samo RDF. Vidjeti: <http://www.schemas-forum.org/>.

⁶Deklaracija RDF sheme za DC nalazi se na web-stranici <http://purl.org/dc/elements/1.1/> pod nazivom *RDF Schema declaration for the Dublin Core Element Set 1.1*.

⁷Baker, Thomas. *A grammar of Dublin Core*. // *D-lib magazine* 6, 10(2000). <http://www.dlib.org/october00/baker/10baker.html>.

dublincore.org/documents/usageguide/" subjekt je povezan s objektom "Diane Hillman" pomoću predikata "creator" koji određuje da je taj objekt stvaratelj.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://dublincore.org/documents/usageguide/">
    <dc:title>Using Dublin Core</dc:title>
    <dc:creator>Diane Hillmann</dc:creator>
    <dc:description>This document is most recent version of the Dublin Core
      usage guide intended as an entry point for users of Dublin Core.</
      dc:description>
    <dc:publisher>Dublin Core Metadata Initiative </dc:publisher>
  </rdf:Description>
</rdf:RDF>
```

Web-stranica opisana je u DC RDF/XML-u. Prvi redak (<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#) u XML-u kaže da se radi o RDF-u definiranom u sintaksi XML-a (xmlns:rdf) prema preporuci Resource Description Framework (RDF) Model and Syntax Specification koja se nalazi na navedenoj web-adresi. Idući redak opisuje RDF shemu i gdje se može naći njezin opis. Radi se o RDF schema declaration for the Dublin Core Element Set 1.1, opisanoj na navedenoj web-adresi. Atribut xmlns povezuje prefiks "dc" s web-adresom DC-a. U trećem se retku navodi izvor koji se opisuje, a to je članak na web-stranici navedene adrese. Od četvrtog do osmog retka koristi se oznaka *description* u kojoj se metapodacima DC-a opisuje izvor. Oznake dc:title, dc:creator, dc:description i dc:publisher pružaju podatke o naslovu, stvaratelju, opisu i nakladniku.

5. Primjeri primjene RDF-a

5.1 UK Online

Na adresi <http://www.ukonline.gov.uk> nalazi se mrežno mjesto Ujedinjenog Kraljevstva (UK Online) za službene informacije raznih vladinih službi, bez obzira na format datoteke. Svi formati datoteka mogu se pregledavati, a svaki je izvor popisan na njihovim stranicama opisan pomoću DC RDF/XML rječnika. UK Online jedinstven je primjer mogućnosti razvoja RDF-a u mrežnom pretraživanju različitih formata dokumenata.

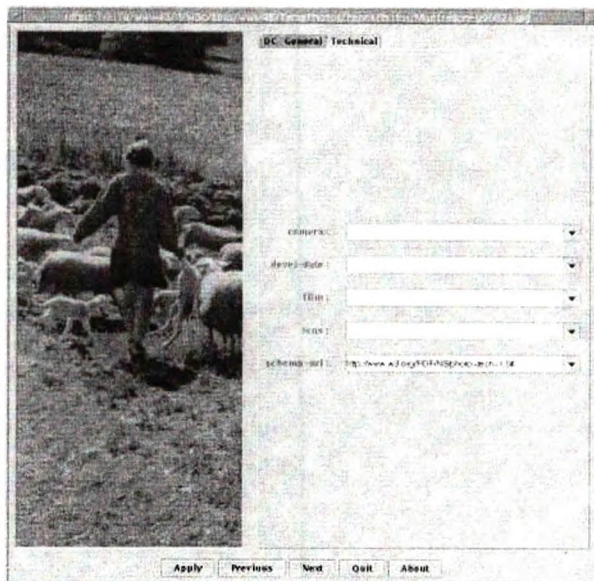
Napredno pretraživanje UK Onlinea nudi pretraživanje po pojedinim riječima iz upita ili po izrazu, naslovu, ključnim riječima i tekstu dokumenta, po posljednjem datumu izmjene dokumenta i sortiranje dokumenata prema relevantnosti ili datumu, sa ili bez sažetka.

5.2 Opisivanje digitalnih fotografija uz pomoć RDF-a

Na mrežnoj adresi <http://www.w3.org/TR/photo-rdf/> nalazi se projekt iza kojeg stoji W3C. Njegov je cilj korištenje RDF-a pri opisivanju digitalnih fotografija kako bi se olakšalo njihovo pronalaženja u mrežnim izvorima. U tu je svrhu napravljen softver i zaseban filter (*plug-in*) za softver za grafičko uređivanje

digitalnih fotografija koji omogućava jednostavno stavljanje metapodataka u RDF-u u i uz sliku u jpeg formatu.

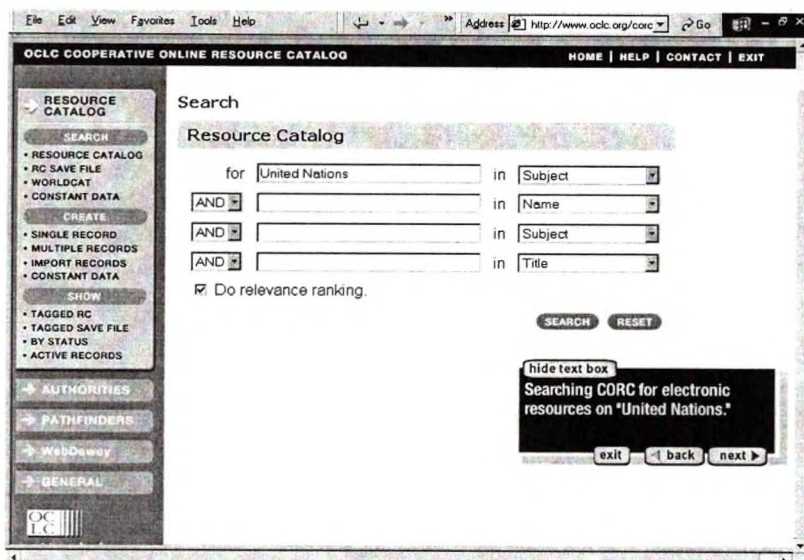
U primjeru na Slici 1 RDF podaci pohranjuju se u datoteku jpeg formata u tzv. blok napomena (*comment blocks*; oznaka je "COM" prema ISO DIS 10918-1). Blok napomena nalazi se unutar digitalnog dokumenta u zasebnim mjestima u koje se inače upisuje format, stvaratelj i još neka od svojstava digitalnog dokumenta. Prema navedenom rješenju ISO-a, upisivanje podataka u blok napomena ne znači i narušavanje autorskih prava ili neke druge strukture ili svojstva digitalnih dokumenata.⁸ U tu se svrhu koriste tri sheme, od kojih je jedna DC.



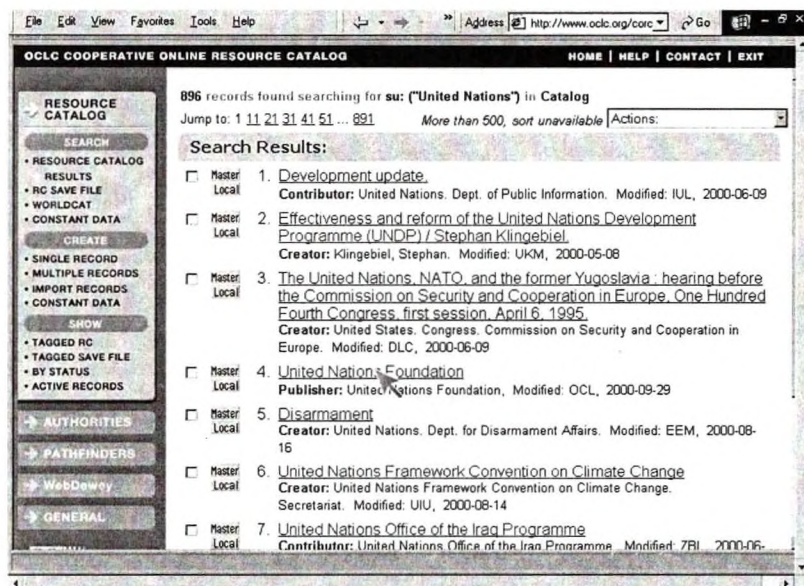
Slika 1. Prikaz *plug-ina* koji koristi softver za grafičko uređivanje za upis RDF shema u tzv. blok napomena digitalne fotografije

5.3 CORC - projekt kooperativne katalogizacije izvora s interneta

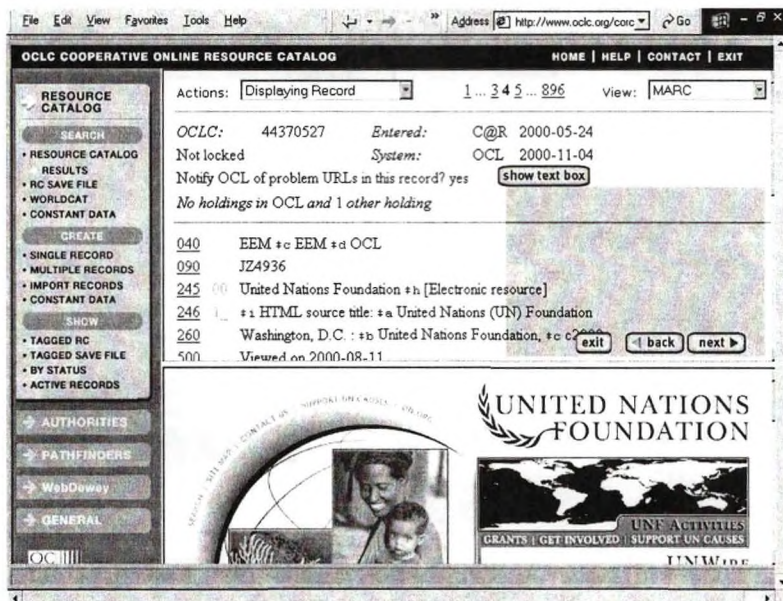
CORC (Cooperative Online Resource Catalog) nalazi se na mrežnoj adresi <http://purl.oclc.org/corc/>. Pokrenuo ga je OCLC kako bi istražio mogućnosti kooperativne katalogizacije mrežnih izvora opisanih u više formata i izradu online automatski ažuriranih mrežnih bibliografija. Na stranici se nalaze alati za automatsku katalogizaciju izvora s interneta, prevođenje formata i sl., uključujući importiranje i eksportiranje zapisa u RDF/XML formatu. Na mrežnoj adresi <http://www.oclc.org/corc/learning/demo/> nalazi se prezentacija CORC-a (slike 2-9).



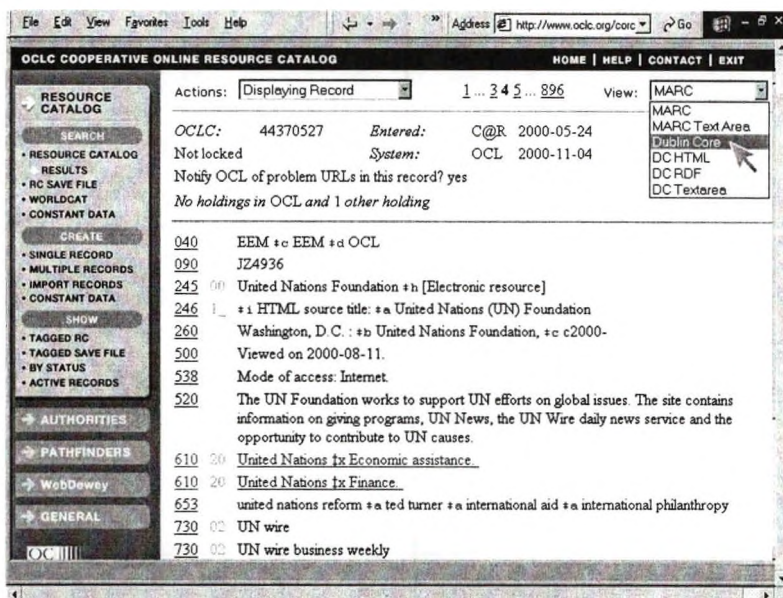
Slika 2. Postavljanje parametara za pretraživanje mrežnih izvora u CORC-u



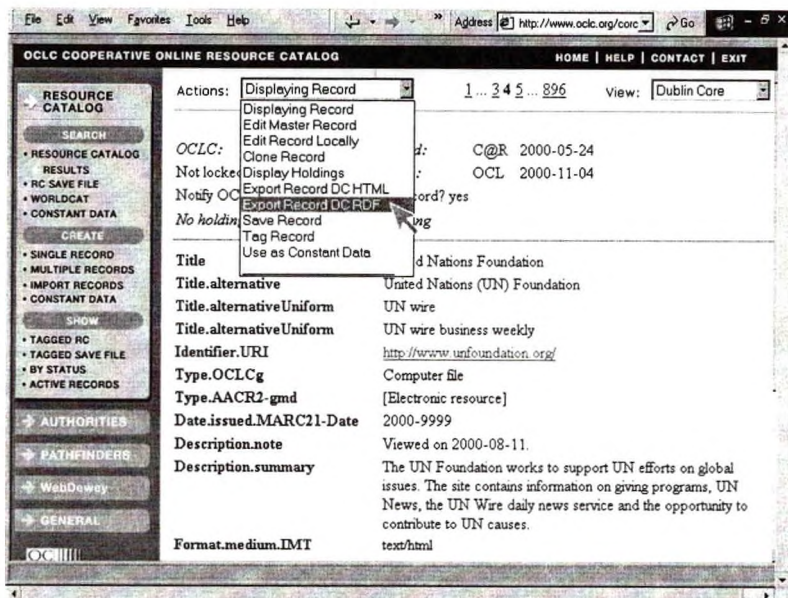
Slika 3. Rezultati pretraživanja na upit "United Nations"



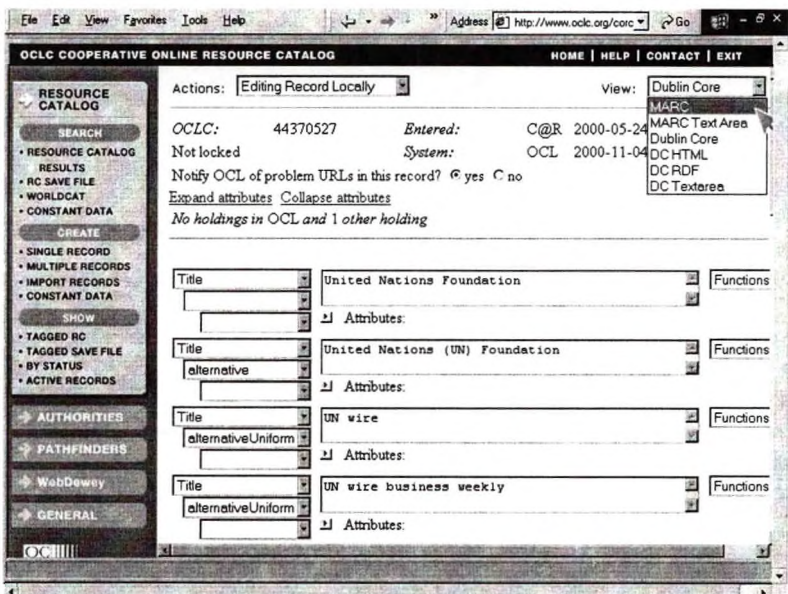
Slika 4. Prikaz mrežnog izvora u MARC formatu



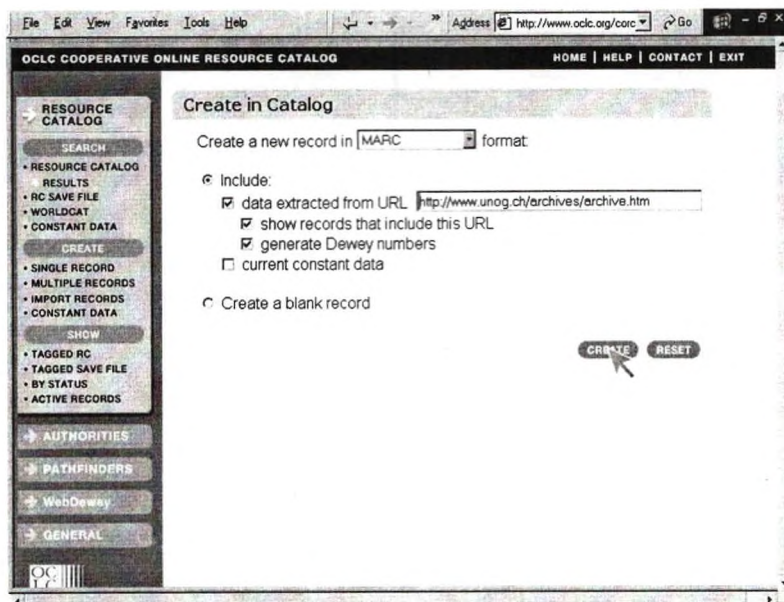
Slika 5. Promjena odabira prikaza iz MARC formata u DC



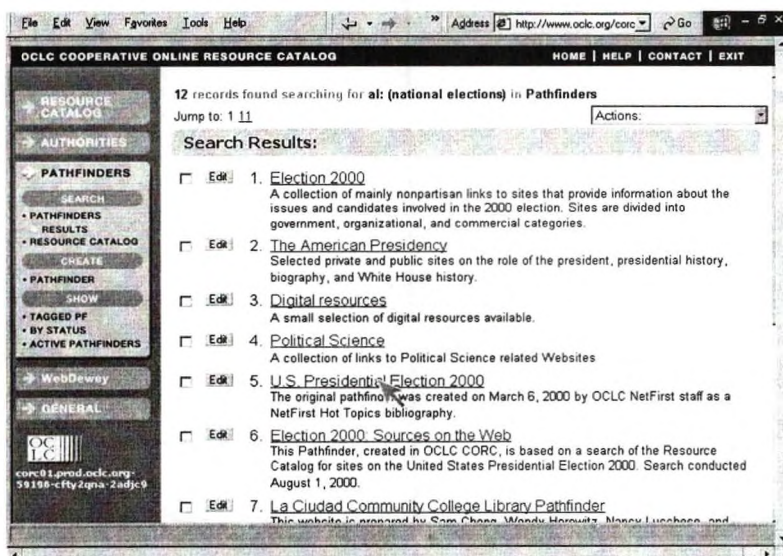
Slika 6. Eksportiranje zapisa u DC RDF-u



Slika 7. Online uređivanje i obrada mrežnog izvora u nekom od formata



Slika 8. Unošenje novog zapisa u CORC upisivanjem URL-a mrežnog izvora; odabran je MARC format



Slika 9. Prikaz mrežne bibliografije u CORC-u

6. Zaključak

XML je omogućio strukturiranje podataka, te razlučivanje podataka koji informaciju predstavljaju, od same informacije. Međutim, različiti sustavi koriste različite podatke za opis (za autora se, primjerice, može rabiti *Creator*, *Author* ili *Written by*). Tu mogućnost da znamo što se smatra pod različitim nazivima koji se odnose na isto, osigurava RDF. RDF model temelji se na tri vrste objekata koji sačinjavaju "rečenicu" za predstavljanje izvora, a sintaksa za prikaz i razmjenu RDF-a je XML.

Mogućnosti weba mogu se u potpunosti ostvariti ako se između sustava mogu prenositi različiti metapodaci. Budući da je RDF taj koji tu komunikaciju i omogućuje, on je koncepcija koja svakako ima budućnost.

LITERATURA

Baker, Thomas. A grammar of Dublin Core. // D-lib magazine 6, 10(2000). <http://www.dlib.org/dlib/october00/baker/10baker.html>.

Champin, Pierre-Antoine. RDF tutorial. <http://www710.univ-lyon1.fr/~champin/rdf-tutorial/>.

Describing and retrieving photos using RDF and HTTP. <http://www.w3.org/TR/photo-rdf/>.

Hunter, Jane; Carl Lagoze. Combining RDF and XML schemas to enhance interoperability between metadata application profiles. <http://archive.dstc.edu.au/RDU/staff/jane-hunter/www10/paper.html> (2001-12-20).

Namespaces in XML. <http://www.w3.org/TR/REC-xml-names/>.

OCLC CORC : web-based metadata creation for describing electronic resources. <http://purl.oclc.org/corc/> (2001-11-15).

RDF schema declaration for the Dublin Core Element Set 1.1. <http://purl.org/dc/elements/1.1/>.

Resource Description Framework (RDF) model and syntax specification. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222> (2001-10-20).

SCHEMAS : Forum for Metadata Schema Designers and Implementers. <http://www.schemas-forum.org/>.

UK Government Information Services. <http://www.ukonline.gov.uk/> (2001-11-15).

World Wide Web Consortium. <http://www.w3.org>.