

# DIGITALNA KNJIŽNICA SLOVENIJE – PORTAL: ZASNOVA IN IZVEDBA

mag. Zoran Krstulović, Matjaž Kragelj, Tine Musek  
Narodna in univerzitetna knjižnica, Turjaška 1, 1000 Ljubljana  
[zoran.krstulovic@nuk.uni-lj.si](mailto:zoran.krstulovic@nuk.uni-lj.si), [matjaz.kragelj@nuk.uni-lj.si](mailto:matjaz.kragelj@nuk.uni-lj.si), [tine.musek@nuk.uni-lj.si](mailto:tine.musek@nuk.uni-lj.si)

## **Povzetek**

Digitalna knjižnica Slovenije ([www.dlib.si](http://www.dlib.si)) je informacijski portal za upravljanje z znanjem. Zasnovan je na strateških ciljnih koncepta Digitalna knjižnica Slovenije – dLib. Osnovno strateško izhodišče je bilo, da je uporabnikom treba prek interneta zagotoviti enotni dostop do knjižničnih katalogov ter zbirk digitaliziranega gradiva, prvotno objavljenega na klasičnih nosilcih. Tehnološke rešitve portala, ki omogočajo dostop do digitalnih vsebin, so zasnovane na implementaciji protokola Open Archive Initiative, avtomatičnemu zajemu podatkov ter metapodatkovnemu modelu Dublin Core. Hkrati portal omogoča tudi dostop do kataloga nacionalne knjižnice prek protokola Z39.50. Tako je uporabnik lahko istočasno preiskuje poljubno število digitalnih oziroma metapodatkovnih zbirk.

## **Abstract**

### **DIGITAL LIBRARY OF SLOVENIA – PORTAL: CONCEPT AND PERFORMANCE**

*Digital library of Slovenia ([www.dlib.si](http://www.dlib.si)) in an informational portal for knowledge management, based on the strategical concepts of the Digital Library of Slovenia – dLib.si. The basic strategic starting point of dLib.si has been to enable the users a simultaneous internet access to library catalogues and collections of materials digitized from classical carriers. Technological solutions of the portal which enable access to digitized contents are based upon the implementation of the protocol Open Archive Initiative, automatic data harvesting and metadata model Dublin Core. At the same time, the portal enables access to the catalogue of the national library using the protocol Z39.50. In this way, a user can search any number of digital or metadata collections.*

## **Ključne besede**

Portal, digitalna knjižnica, Dublin core, protokoli, Z39.50, Open Archive Initiative (OAI).

## **1. UVOD**

Med direktivami, pobudami in drugimi dokumenti Evropske unije in Republike Slovenije najdemo dovolj takih, ki obravnavajo digitalizacijo, digitalno dediščino, "digitalni srednji vek", e-repozitorij in druge elemente knjižnice v internetnem okolju ter opredeljujejo načela, na katerih sloni strategija razvoja *Digitalne knjižnice Slovenije – dLib.si*. V mednarodni skupnosti je soglasje glede metodologije uresničevanja digitalne knjižnice šibko, še manj pa je mednarodno veljavnih standardov, posebej na področju trajnega ohranjanja digitalnega gradiva. Digitalna knjižnica zahteva infrastrukturo, ki zagotavlja njeno delovanje v dolgem časovnem obdobju ter nenazadnje potrebuje jedro strokovnjakov, usposobljenih za implementacijo novosti, ki jih prinaša tehnološki razvoj ter zahteva permanentno vzdrževanje digitalnih arhivov. Na drugi strani pripravljenost knjižničarjev na spremembe, ki jih prinaša sprememba paradigme knjižničarstva, šele zori. Iz varuhov kulturne dediščine, skrbnikov katalogov ter informacijskih strokovnjakov se knjižničarji spreminjajo v knjižničarje-informatike, ki morajo poleg klasičnih bibliotekarskih znanj in veščin obvladovati in obvladati metodologije izgradnje informacijskih sistemov in kreiranja digitalnih zbirk, organizacijo in diseminacijo znanja prek informacijskih portalov ipd. Zaradi vsega

omenjenega smo si postavili za cilj, da *Digitalna knjižnica Slovenije* postane žarišče sprememb ter infrastruktura nove paradigme knjižničarstva.

Med najpomembnejšimi dejavniki, ki bodo v prihodnje vplivali na razvoj evropskih znanstvenih knjižnic, bo omrežje (network) postalo primarni element v informacijski verigi. Večina gradiva v klasični obliki, ki ga danes knjižnice hranijo na policah, bo digitalizirana in dostopna prek omrežij. Eden izmed najpomembnejših izzivov knjižnic v prihodnosti bo arhiviranje in ohranitev digitalnega oziroma elektronskega gradiva za prihodnje rodove. Omrežja bodo omogočila večji obseg sodelovanja in delitve dela med knjižnicami na regionalni, nacionalni in mednarodni ravni, zato se bo pokazala potreba po opredelitvi standardov, ki bodo olajšali elektronsko izmenjavo informacij in njihovo iskanje. Pri posredovanju dokumentov se bodo knjižnice zagotovo soočale s konkurenco komercialnih ponudnikov informacijskih virov.

Danes je gradnja digitalnih zbirk ena izmed pomembnih dejavnosti knjižnic, vse bolj pa postaja tudi osrednja naloga, še zlasti nacionalnih knjižnic. Zato smo v strateškem načrtu slovenske osrednje knjižnice Narodne in univerzitetne knjižnice (NUK) zapisali kot vodilo, da je prav nacionalna knjižnica "most med preteklo in sedanjo duhovno ustvarjalnostjo slovenskega naroda in vrata v svet informacij in znanja".<sup>1</sup> Kot eno izmed osrednjih izhodišč strateškega načrta je zapisana teza, po kateri "Slovenija bo razvojno sposobna enakopravno tekmovati znotraj EU, če bo vzpostavila učinkovito digitalno knjižnico!"<sup>2</sup> Dejstvo je, da je danes prek medmrežja v celotnem besedilu v slovenščini dostopnih le malo tehtnih dokumentov oz. vsebin. Za razvoj znanosti in raziskovanja, za izobraževanje in strokovno delo v vseh dejavnostih pa je pomembno, da bodo čim prej vsi pomembni dokumenti, publikacije, raziskovani dosežki itd., domači in tuji, dostopni v elektronski obliki na delovnem oz. študijskem mestu. Uresničitev teh strateških smernic bo bistveni sestavni del oblikovanja slovenske digitalne knjižnice.<sup>3</sup>

## 2. URESNIČITEV STRATEŠKIH CILJEV: ZASNOVA PORTALA

Razlog za izgradnjo portala je nadgradnja uresničevanja strateških ciljev, zapisanih v izvedbenem dokumentu strateškega načrta NUK *Digitalna knjižnica Slovenije – dLib*<sup>4</sup>, in sicer: prek interneta zagotoviti dostopnost do digitaliziranega gradiva, objavljenega na klasičnih nosilcih, ter vzpostavitev portala digitalne knjižnice Slovenije.

Razdrobljenost podatkovnih digitalnih zbirk in katalogov, ki so prek različnih, bolj ali manj vzdrževanih spletnih vmesnikov dosegljivi uporabnikom, zahteva na več ravneh enotnost in transparentnost. Dejstvo, da je spletna mesta potrebno vzdrževati, izpopolnjevati ter (oblikovno) spreminjati, je zgolj eden od vidikov, ki upravičuje snovanje izgradnje enotnega spletnega vmesnika za pregled digitalnih vsebin. Nepravilnosti, ki se lahko pojavijo v "življenjskem obdobju" aplikacije oziroma spletnega portala, določajo njihov obstoj, razvoj in rast same spletne vsebine, so med drugim dosegljivost, izvoz in uporaba podatkov – digitalnih vsebin. S tega vidika portal digitalne knjižnice Slovenije sledi metodologiji portala evropske

---

<sup>1</sup> Strateški načrt NUK 2004-2008. Poslanstvo, vizija, vrednote in cilji. Ljubljana, Narodna in univerzitetna knjižnica, 2004, str. 2.

<sup>2</sup> Id.

<sup>3</sup> Id.

<sup>4</sup> Krstulović, Zoran, Šetinc, Lenart: Digitalna knjižnica Slovenije – dLib.si, Zbornik posvetovanja Dnevi slovenske informatike 2005, Portorož, 2005, str. 683-689.

digitalne knjižnice, *The European Library* – TEL<sup>5</sup>, katere zbirke so precej večje – tako po številu, kot tudi po količini zapisov v njih, sama zamisel shranjevanja podatkov ter format zapisa le-teh pa sta primerljiva.

## 2.1 Trenutno dostopne zbirke na portalu

Spletni portal *dLib* na enem mestu omogoča dostop do različnih digitalnih vsebin s pomočjo metapodatkov, ki jih shranjuje v svoji podatkovni bazi. Trenutno *dLib* omogoča dostop do Kataloga NUK in petih digitalnih oziroma metapodatkovnih zbirk:

- Slovenska bibliografija,
- Slovenska retrospektivna bibliografija člankov (s polnimi besedili v digitaliziranih revijah *Ljubljanski zvon*, *Novi akordi* in *Nova muzika*),
- Zbirka portretov znanih Slovencev in Slovenk,
- Zbirka zvočnih posnetkov slovenske glasbe,
- Zbirka Zemljevidov slovenskega ozemlja.

## 2.2 Pridobivanje zbirk v potrebnem formatu

Ustanova, ki vzdržuje digitalno zbirko in želi zagotoviti njeno dostopnost prek portala *dLib*, mora poskrbeti za izvoz in konverzijo (prek javno dostopnih konverzijskih tabel) bibliografskih zapisov zbirke v metapodatkovni standard Dublin core<sup>6</sup> (dalje DC). DC je metapodatkovni standard za iskanje in poizvedovanje po multimedijских informacijah oz. internetnih virih. Gre za skupno metapodatkovno jedro, ki je prenosljivo in razumljivo v posameznih specifičnih okoljih. Elementi tega modela so polja, katerih prisotnost, vrstni red ter pogostnost so poljubni (npr. kot del XML zapisa). Dublin core je sestavljen iz petnajst osnovnih polj, in sicer:

- Naslov (Title)
- Avtor (Creator)
- Vsebina (Subject)
- Opis (Description)
- Založnik (Publisher)
- Sodelavec(-ci) (Contributor)
- Datum (Date)
- Zvrst (Type)
- Format (Format)
- Identifikator (Identifier)
- Vir (Source)
- Jezik (Language)
- Odnos do drugih del (Relation)
- Pokrivanje (Coverage)
- Pravice (Rights).

NASLOV: Za vasjo, valček

AVTOR: Viktor Pokorn

VSEBINA: narodnozabavna glasba, slovenija, ok. 1948, pihalna godba, valčki, polke

ZALOŽNIK: Zagreb, Jugoton

---

<sup>5</sup> [www.theeuropeanlibrary.org](http://www.theeuropeanlibrary.org).

<sup>6</sup> Dublin Core Metadata Initiative, <http://dublincore.org>.

SODELAVEC: Kmečka godba iz Ljubljane  
DATUM: ok. 1948  
ZVRST: zvočni posnetki, glasbeni, monografska publikacija  
JEZIK: slovenski  
IDENTIFIKATOR: povezava do gradiva

**Tabela 1: Primer metapodatkovnega zapisa v DC formatu**

V podatkovnih zbirkah hranimo torej metapodatkovne opise drugih podatkovnih zbirk, ki so lahko shranjeni v poljubni bazi podatkov, na poljubni platformi. Vsebine in strukture teh baz ter formata njenih elementov ne vidimo. Vse naše védenje o teh zbirkah je omejeno na metapodatkovni opis njenih elementov. Zastavlja se vprašanje, kako zagotoviti koherentnost naših podatkov ter podatkov v podatkovni zbirki lastnika zbirke.

## **2.3 Načini zbiranja podatkov**

### **2.3.1 Open Archive Initiative (OAI)**

Bistvo odprtega arhiva (po zgledu odprte kode) je omogočiti do vsebin, ki so dostopne prek spleta. Te vsebine so ponavadi shranjene v interoperabilnih repozitorijih, ki zbirajo, arhivirajo in ponujajo metapodatke. Potreba po takem načinu izmenjave podatkov je nastala v elektronskem založništvu, kjer iz dneva v dan rastejo zahteve po preprostem in interoperabilnem dostopu do podatkov, ki so ponavadi shranjeni v heterogenih repozitorijih (različne platforme, različne baze). Tako je nastala pobuda odprtega arhiva, po angleško *Open Archive Initiative* ali skrajšano *OAI*.

OAI razvija in uveljavlja preprosto delovno okolje, znotraj katerega uporablja standarde, ki so spočetka lajšali delo v elektronskem založništvu, danes pa se OAI uporablja tudi za dostop do mnogih drugih, prek spleta dostopnih digitalnih vsebin. "OAI razvija in uveljavlja interoperabilne standarde, ki naj bi pospešili in poenostavili učinkovitost razbitih virov podatkov."<sup>7</sup>

Tako kot pri odprti kodi je širok krog sodelujočih pri razvoju za večino ugoden. Enako velja tudi za odprti arhiv. OAI resda izvira iz Združenih držav Amerike, kjer je tudi sedež Izvršnega odbora, največji uspehi pri uveljavitvi in uporabi standarda pa se kažejo prav v Evropi. Prav tako ni nobene potrebe po tem, da bi peščica ljudi držala v svojih rokah kontrolo nad standardom, posebej sedaj, ko je v veljavi že druga verzija protokola OAI, ki je natančno določen in stabilen za uporabo. Zato je na skupščini OAI v postopku odločitev, da bi organizacija predala kontrolo nad razvojem protokola v roke združenja za standardizacijo ISO. Do sedaj je to delo opravljal odbor za usmerjanje in tehnično direktivo razvoja standardov znotraj organizacije OAI.

Vedno več združb ugotavlja uporabnosti in prednosti protokola OAI, saj se količina digitalnih vsebin spletu iz dneva v dan povečuje, s tem pa se povečuje tudi število uporabnikov različnih podatkovnih repozitorijev. S pomočjo OAI protokola se dostopnost podatkov poenostavi in jih je mogoče uporabiti tudi v druge namene, za kakršne so bili prvotno predvideni, kar zmanjšuje stroške izdelave novih komercialnih vsebin. Poleg tega možnost dostopa do več repozitorijev prinaša razvoj novih storitev, bolj prijaznih do uporabnika.

### **2.3.2 OAI protokol za avtomatično zajemanje metapodatkov (OAI-PMH)**

---

<sup>7</sup> Open Archives Initiative. Mission Statement, <http://www.openarchives.org/organization/index.html>.

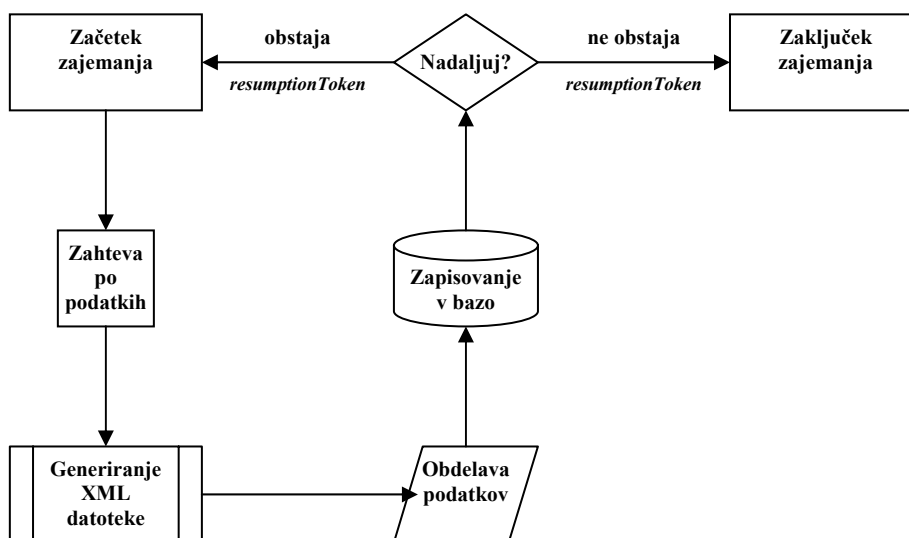
OAI protokol za avtomatično zajemanje metapodatkov (dalje OAI-PMH) natančno določa mehanizme za zajem metapodatkovnih zapisov iz repozitorijev in ponuja preprosto in tehnično nezahtevno možnost vsem ponudnikom, da svoje podatke posredujejo storitvam, ki so zasnovane na uveljavljenih protokolih HTTP in XML. Metapodatki sami so lahko v kakršnemkoli formatu. Končna verzija oblike posredovanja podatkov je odvisna od dogovora med ponudnikom in prejemnikom, čeprav neuradno velja, da naj bi bili podatki v formatu DC. Na ta način lahko nastajajo veliki podatkovni repozitoriji ali agregati, ki ponujajo uporabniku dostop do vseh podatkov z enega mesta. Pomembno je poudariti, da povezave med metapodatki in dejansko vsebino OAI protokol ne določa, kar pomeni, da OAI-PMH ne išče digitalne vsebine v prenesenih zapisih, ampak le ponuja možnost zbiranja metapodatkov.

Sam protokol veliko obeta, nekateri celo pričakujejo, da bo postal eden od sestavnih delov infrastrukture svetovnega spleta, saj ga, podobno kot protokol HTTP, opredeljuje preprostost, zanesljivost in predvsem interoperabilnost. Čeprav je sam OAI-PMH tehnično nezahteven, izgradnja koherentne in uporabniku prijazne storitve še vedno ostaja zapletena. Cilji in prihodnost protokola vidijo predvsem kot nov vzorec za komunikacijo med izobraževalnimi ustanovami in storitvami za posredovanje znanja in/ali digitalnih vsebin.

### 2.3.3 Aplikacija za avtomatični zajem podatkov (harvester)

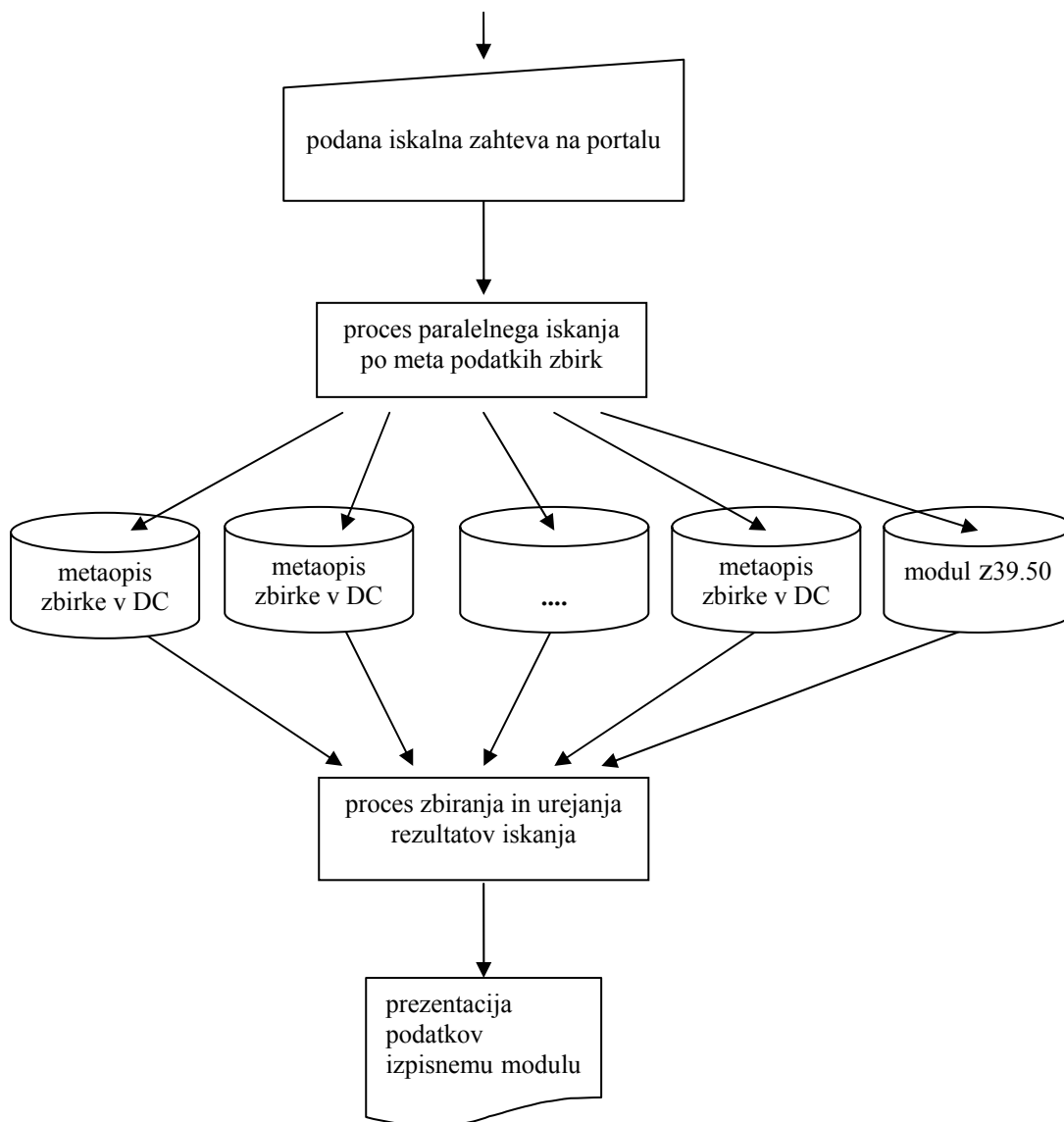
Sama gradnja repozitorija *dLib* je zahtevala poenotenje formatov, v katerih so zapisani metapodatki. Prav tu OAI v kombinaciji z DC ponuja enostavno, hitro in graditelju repozitorija prijazno okolje, samemu oskrbniku/ponudniku pa nižjo ceno obratovanja. Za bibliografsko bazo, kakršno jo uporablja *dLib*, smo uporabili XML shemo za OAI, kot jo predpisuje *The European Library*.

Potek zajema podatkov se začne z zahtevo na OAI-PMH strežnik. Ta vrne rezultat v vnaprej določeni obliki in shemi v XML formatu. Ker so baze lahko ogromne in bi izgradnja in prenašanje velikih XML datotek lahko predolgo trajalo, se prenos baze razbije v poljubno število nizov, vsak niz pa se konča s poljem (*resumptionToken*), s pomočjo katerega strežniku pošljemo zahtevo za naslednji niz zapisov. Če polja ni, aplikacija za zajem predvideva, da je prišla do konca baze in preneha z zbiranjem podatkov. Shematično delovanje lahko vidimo na spodnji sliki.



Slika 1: Shema zajema podatkov

Zajem poteka prek protokola HTTP, kar pomeni, da je prenos podatkov počasen. Zato je uporaba OAI v *dLibu* smiselna le pri zaključenih zbirkah/bazah podatkov, kjer je pogostnost spreminjanja podatkov nizka, zajem pa se lahko dogaja v nočnih urah, ko je uporaba portala majhna. Za pregledovanje "žive baze" pa uporabljamo druge načine dostopanja do podatkov, npr. protokol Z39.50.



Slika 2: Modul za iskanje po dLib-ovih (digitaliziranih) podatkovnih zbirkah

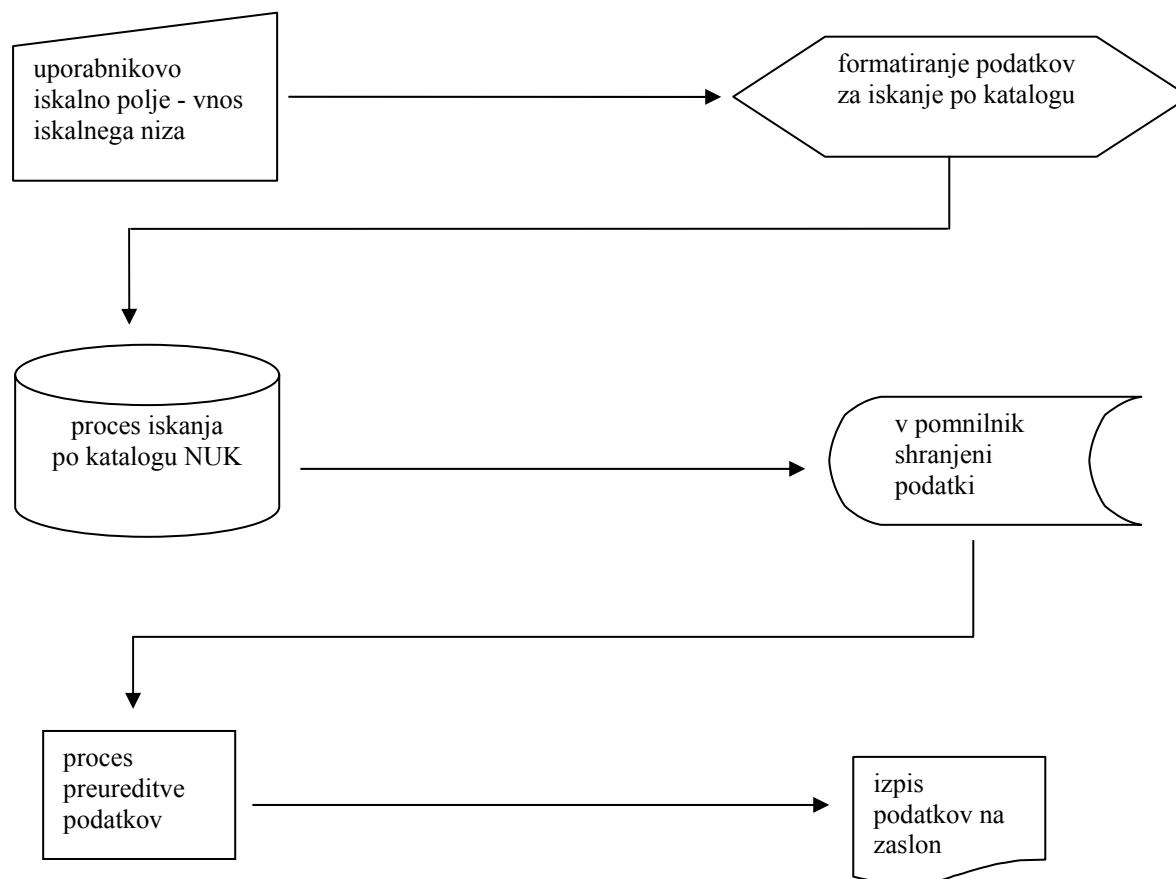
### 2.3.4 Z39.50 in dostop do lokalnega kataloga

Poleg iskanja po metapodatkih digitalnih zbirk, *dLib* omogoča tudi iskanje po katalogu NUK. Načelo delovanja je drugačno kot pri iskanju po digitalnih zbirkah, saj poteka prek strežnika na Institutu informacijskih znanosti (IZUM) v Mariboru, ki podpira protokol z39.50<sup>8</sup>, ter priklopa na lokalno bazo – katalog NUK. Podatki podatkovne baze kataloga so v formatu COMARC. Poskus gradnje lokalne zbirke metapodatkov vsebin kataloga ni najboljša rešitev,

<sup>8</sup> About the Z39.50 Gateway, <http://www.loc.gov/z3950/gateway.html#about>

predvsem zaradi nenehnega poseganja v katalog s strani katalogizatorjev gradiva (problem koherentnosti) in velikanske količine gradiva v njem.

Iskanje prek dLib-a aktivira povezavo na strežnik *z3950.izum.si* ter prijavo na bazo NUK. Izvede se iskanje na strežniku, rezultat se vrne dLib-u, kjer se opravi prevod iz formata COMARC v DC ter ponudi uporabniku.



Slika 3: Modul za iskanje po lokalnem katalogu z uporabo protokola Z39.50

### 3. SKLEP

Knjižnice bodo temelj družbe znanja, v njih shranjeno znanje pa bo moralo postati čim lažje in široko dostopno, pri čemer lokacija uporabnika ne bo ovira pri dostopu do informacij. Paradigma interneta in zahteva po neomejeni dostopnosti do informacij oz. znanja vedno bolj prežemata in preoblikujeta knjižničarsko stroko samo. Zaznavamo lahko tudi spremembo paradigme knjižničnega poslovanja, ki se od nekdanje prevladujoče manipulacije s posameznimi kosi knjižničnega gradiva (nabava, bibliografska obdelava, skladiščenje na policah, izposoja) spreminja v smeri organizacije podatkov/znanja (konverzije, avtomatska obdelava, elektronski servisi, zbirke polnih besedil). Danes integrirani knjižnični sistem kot orodje za dostop do informacij ni več zadostna podpora niti avtomatiziranemu knjižničnemu poslovanju niti uporabnikom zlasti zato, ker ne omogoča učinkovitega in enotnega dostopa do internetnih vsebin (licenčnih virov, podatkovnih zbirk, elektronskih časopisov) in ne zadovoljuje potreb

knjižnic po upravljanju z drugimi elektronskimi viri (digitaliziranimi zbirkami), ki nastajajo v knjižnicah. Zato sodobna knjižnica za podporo knjižničnemu poslovanju potrebuje nova orodja. Poleg integriranega knjižničnega sistema, ki knjižničarjem omogoča bibliografsko obdelavo, izposajo gradiva, statistike ipd., uporabniku pa poizvedovanje po knjižničnem katalogu, potrebuje tudi sistem za upravljanje z digitaliziranimi podatkovnimi zbirkami ter spletni portal za upravljanje z znanjem, ki združuje uporabniške servise in poslovne funkcije knjižnice. Pričakujemo, da bo v prihodnosti modularnost imela prednost pred monolitno integriranostjo knjižničnih sistemov. Modularnost bo zahtevala tehnično interoperabilnost, ki jo bo zagotavljala široka paleta mednarodnih, nacionalnih in celo industrijskih standardov. Integrirani sistemi bodo nadomeščeni z modularnimi sistemi tj. sistemi, ki se bodo razvili iz paradigme odprtosti in neomejenega dostopa do interneta in bodo omogočali knjižnici kombinirati najboljše značilnosti različnih izdelkov<sup>9</sup>.

Paradigma knjižničarstva se spreminja v skladu s predpostavko, da bo ravno sposobnost za upravljanje z elektronskimi viri naslednji veliki korak v avtomatizaciji knjižničnega poslovanja<sup>10</sup>. Trenutno smo še vedno zaprti v okvire, ki jih določa potek dela znotraj tradicionalnega knjižničnega sistema, se pravi, da se večina podatkov v integrirani sistem vnaša ročno. Novi sistemi za upravljanje z digitaliziranimi podatkovnimi zbirkami pa bodo omogočali avtomatično procesiranje različnih podatkov.

#### 4. VIRI IN LITERATURA

- [1] About the Z39.50 Gateway, <http://www.loc.gov/z3950/gateway.html#about>  
Dublin Core Metadata Initiative, <http://dublincore.org>.
- [2] HAKALA, Juha: There and back again: from integrated to modular library systems, Helsinki University Library Bulletin, 2003, str. 11-14.
- [3] KRSTULOVIC, Zoran, ŠETINC, Lenart: Digitalna knjižnica Slovenije – dLib.si, Zbornik posvetovanja Dnevi slovenske informatike 2005, Portorož, 2005, str. 683-689.
- [4] Open Archives Initiative. Mission Statement, <http://www.openarchives.org/organization/index.html>.
- [5] Strateški načrt NUK 2004-2008. Poslanstvo, vizija, vrednote in cilji. Ljubljana, Narodna in univerzitetna knjižnica, 2004, str. 2.
- [6] [www.theeuropeanlibrary.org](http://www.theeuropeanlibrary.org).

---

<sup>9</sup> Hakala, J.: There and back again: from integrated to modular library systems, v: Helsinki University Library bulletin 2003, str. 11.

<sup>10</sup> Ib., str. 12.