

SREČANJE UPORABNIKOV SISTEMA CBS 2014

RAZVOJ BIBLIOGRAFSKIH FORMATOV IN KATALOGIZACIJSKIH PRAVIL

Evropska podružnica OCLC (Online Computer Library Center, www.oclc.org) je septembra 2014 organizirala dvodnevno srečanje uporabnikov sistema CBS (Centraal Bibliografisch Systeem) z naslovom *The evolution of library data formats and cataloguing conventions*. Srečanje je potekalo v idiličnem univerzitetnem mestu Leiden na Nizozemskem. Udeležili so se ga predstavniki Nemčije, Francije in Švice, predstavniki IZUM-a pa kot opazovalci. CBS je knjižnični sistem, ki ga uporabljajo knjižnice na Nizozemskem, v Nemčiji, Franciji in Švici; uporabljajo ga tudi v Nacionalni knjižnici Avstralije. Sistem je osnovala nizozemska PICA, ki se je leta 1997 združila z OCLC-jem v OCLC PICA (en.wikipedia.org/wiki/OCLC_PICA), od leta 2007 pa je v celoti integrirana v OCLC.

NOVOSTI V OCLC

Ted Fons (izvršni direktor, OCLC Data Services and WorldCat Quality Management) je v predstavitvi z naslovom *The evolution of WorldCat* pojasnil nov koncept knjižničnih podatkov, ki bo omogočal vsebinsko povezovanje podatkov in konceptov o različnih tipih objektov (npr. o osebah, mestih ali celo bolj abstraktnih stvareh, kot so teme, dogodki ...) in ne samo povezovanje dokumentov na spletu; s tem bi se znatno izboljšala uporabniška izkušnja pri iskanju in predstavitvi podatkov iz knjižničnih katalogov. Ljudje namreč na spletu nočejo več samo prikaza seznamov, npr. o knjižnem gradivu in avtorjih, ampak tudi kontekst in podrobnosti o neki osebi, organizaciji, sestanku itd. Uporabniki pravzaprav iščejo delo po modelu FRBR (www.ifla.org/publications/functional-requirements-for-bibliographic-records). V bibliografskem zapisu, ki ustreza pojavnim oblikam po tem modelu, pa so opisane podrobnosti o določenem delu. Dosedanje bibliografske in normativne zapise bo nadomestil model entitet, ki predstavljajo vsebinsko ločene dele zapisa s podatki o delu, osebah, organizacijah, temah, mestih in dogodkih. Zato se bo spremenil potek dela v knjižnicah:

- Odkrivanje podatkov (angl. *discovery*) bo postalo bolj privlačno in uporabno, med entitetami se bodo vzpostavile povezave (angl. *entity based linking*), kar zahteva, da bo za osebe, dela in druge entitete treba

kreirati normativne oblike. Vzpostaviti in kreirati bo treba nove vrste normativnih datotek.

- Katalogizacija bo doživela največje spremembe, saj bo postala katalogizacija tipa "označi in klikni" (angl. *point and click*). Namesto bibliografskih zapisov bomo vodili in povezovali entitete, zato se bodo za vse entitete vzpostavili identifikatorji, katalogizacija pa bo potekala z združevanjem posameznih entitet v en zapis. Kaskadno dopolnjevanje (ažuriranje samo posameznih entitet in ne celega bibliografskega zapisa) bo povečalo kvaliteto metapodatkov in omogočilo preprostejše vzdrževanje. Upravljanje metapodatkov oz. zapisov bo programsko, temeljilo bo na zajemanju podatkov o entitetah iz različnih informacijskih virov in baz podatkov. Knjižnice bodo uporabljale predvsem katalogizacijska pravila RDA (Resource Description and Access, en.wikipedia.org/wiki/Resource_Description_and_Access).
- Prikaz podatkov na spletu bo bogatejši in bo uporabnike dejansko pripeljal do informacij, ki so dostopne na osnovi povezanih podatkov.

OCLC je aprila 2014 objavil bazo podatkov WorldCat Works (www.oclc.org/data.en.html) s 197 milijoni zapisov za dela s povezavami do pojavnih oblik del (bibliografskih zapisov) v bazi podatkov WorldCat ter povezavami z zapisi iz baz DDC, VIAF (viaf.org), FAST (www.oclc.org/research/themes/data-science/fast.html), LCSH (authorities.loc.gov) in iz drugih normativnih baz podatkov. Model entitet ni enak modelu FRBR in za zdaj predstavlja lastno rešitev za WorldCat.org. Omenil je še, da ima Bibframe (www.loc.gov/bibframe) nekoliko drugačen pristop, ker izhaja iz perspektive katalogizacije, schema.org (schema.org), ki je osnova za model entitet v WorldCat.org, pa iz spletne perspektive.

Axel Kaschte je v predavanju *Metadata globalization and the European cataloguer* predstavil prehod programske opreme na oblačne storitve ter licenčnih pogodb na naročnino. Vse to zahteva nov način dela, zato bodo knjižnice v sistemu CBS postopno prehajale na OCLC-jevo platformo WorldShare. Prehod otežuje raznolikost knjižnic v sistemu CBS; francoske knjižnice uporabljajo format UNIMARC, druge pa format MARC 21, vse pa za nabor znakov uporabljajo razširjeni UNICODE.

Martin van Muijen je predstavil skupino Metadata Advisory Group (MAG); njen glavni namen je priprava navodil za evolucijo metapodatkov. OCLC množično vključuje mednarodne baze podatkov v WorldCat, ki vsebuje vzporedne bibliografske zapise glede na jezik katalogizacije. Za zbirke elektronskih virov uporabljajo bazo znanja (angl. *knowledge base*). Model entitet so začeli uporabljati za knjige in revije, za članke pa še ne. Načrtujejo, da bo vzporedno okolje bibliografskih zapisov in modela entitet v uporabi še zelo dolgo oz. dokler bo potrebno.

NOVOSTI V SISTEMU CBS

Bemal Rajapatirana (OCLC) je v predavanju z naslovom *The impact on CBS of the UKB transition* predstavila prehod nizozemskega nacionalnega kataloga (GGC), pri katerem sodelujejo nizozemske univerzitetne knjižnice in nacionalna knjižnica UKB (*Koninklijke Bibliotheek*), v bazo podatkov WorldCat. Po prehodu bodo knjižničarji katalogizirali neposredno v bazo podatkov WorldCat in ne več v nacionalni katalog. Tudi za vse preostale postopke bodo uporabljali OCLC-jeve servise, kot sta npr. administracija v bazi znanja in medknjižnična izposoja v sistemu WorldShare ILL. Pri katalogizaciji bodo uporabljali normativne datoteke Kongresne knjižnice, za e-vire pa bodo zalogo kreirali avtomatsko. Posebne zbirke bodo vključevali kasneje. CBS bo pripravil konverzijo bibliografskih podatkov iz formata PICA (www.gbv.de/wikis/cls/PICA-Format) v format MARC 21, sledil bo paketni vpis v bazo WorldCat. Za prikaz zaloge lokalnih knjižnic bodo v WorldCat pošiljali akronime knjižnic. Pri katalogizaciji v bazi podatkov WorldCat bodo katalogizacijska pravila uporabljali selektivno glede na vrsto publikacije. Za zapise nacionalnega kataloga bodo še naprej uporabljali nacionalna pravila, za e-vire pa pravila RDA. Povedala je tudi, da bo konzorcij GGC obstajal, dokler bo to potrebno.

Uwe Risch (HeBIS) je v predavanju z naslovom *The progress of the CIB project* predstavil nemški projekt CIB (*Cloudbasierte Infrastruktur für Bibliotheksdaten*, www.projekt-cib.de/wordpress), ki ga vodi HeBIS (Hessisches Bibliotheksinformationssystem), sodelujejo pa še BVB (Bibliotheksverbund Bayern), KOBV (Kooperativer Bibliotheksverbund Berlin-Brandenburg) in DFG (Deutsche Forschungsgemeinschaft). Pripravljajo katalogizacijski sistem v oblaku, ki mora biti po pravilih evropske zakonodaje lociran v Nemčiji ali v drugi državi znotraj Evropske unije. Namesto lokalnih pravil bodo začeli uporabljati mednarodne standarde za metapodatke in jezikovno neodvisne kode. Knjižnicam ne bodo predpisali knjižničnega sistema, pričakujejo pa, da bodo uporabljale OCLC-jeve ali Exlibrisove servise.

Tobias Viegner je predstavil prenos kataloga Švicarske

nacionalne knjižnice in univerzitetnih knjižnic SwissBib (www.swissbib.ch) v WorldCat. Za prenos zapisov so pripravili model glavnega zapisa (angl. *master record*), pri katerem je v enem zapisu združenih več različnih zapisov za isto publikacijo. Prenos bo paketni in brez ročne katalogizacije. Ko je zapis enkrat vpisan v bazo podatkov, ga več ni mogoče spreminjati, ampak se samo dopolnjuje z novimi podatki. Vsaka sprememba v lokalni bazi se vidi tudi v vzajemni bazi. V WorldCat pošiljajo samo glavne zapise. Model glavnega zapisa uporabljajo tudi v drugih servisi; tudi v druge vire in baze podatkov pošiljajo zapise na osnovi glavnega zapisa. Pri združevanju zapisov uporabljajo strogo in prilagodljivo ujemanje podatkov (angl. *strict and flexible matching*). Zaradi nenehnega pregrupiranja (angl. *reclustering*) zapisi nimajo stalnega ID-ja.

Reinhard Alrtenhoener (DNB) je v predavanju *GND and ZDB in the globalized infrastructure remarks from the DNB perspective* predstavil nemške normativne baze podatkov in stanje v nemških knjižnicah. GND (Gemeinsame Normdatei; www.dnb.de/DE/Standardisierung/GND/gnd_node.html) je normativna baza osebnih in korporativnih imen, sestankov, predmetnih oznak in del z okrog 10 milijoni zapisov. Ta baza podatkov je pomemben vir za mednarodni agregator normativnih podatkov VIAF in za mednarodni standardni identifikator imen ISNI (www.isni.org). ZDB (Zeitschriftendatenbank; www.zeitschriftendatenbank.de) je baza podatkov o serijskih publikacijah z okrog 1,6 milijona zapisov. Nemške knjižnice uporabljajo format MAB (www.dnb.de/EN/Standardisierung/Format/MAB/mab_node.html; njegovo vzdrževanje so ukinili leta 2013) ali format PICA, izmenjava med njima pa poteka v formatu MARC 21. Izmenjava podatkov bo avtomatizirana, ročna bo le končna verifikacija. Serijske publikacije bodo katalogizirali neposredno v bazo podatkov WorldCat ali z ExLibrisovim programom Alma. Po prehodu bodo v bazi podatkov ZDB samo še naslovi serijskih publikacij in zaloga. Zaenkrat še ne vedo, kakšna bo povezava med Registrom ISSN in bazo podatkov ZDB. Zapisov ne združujejo in ne izločajo podvojenih zapisov (duplikatov), ampak uporabljajo grupiranje zapisov (angl. *clustering*). Na koncu je omenil, da postaja WorldCat vse pomembnejši.

Martin van Muijen (OCLC) je v zaključnem predavanju *CBS roadmap* pripravil pregled projektov zadnjega leta. Model FRBR uporabljajo samo članice CBS na Nizozemskem, izboljšali so metodo združevanja zapisov v entitete FRBR, CBS-ov model glavnega zapisa uporablja samo SwissBib, povezava med sistemoma CBS-ILL in VDX za medknjižnično izposajo uporabljajo samo knjižnice na Nizozemskem, v spletnem vmesniku so vključili pomoč. Sinhronizacije lokalne zaloge z bazo podatkov WorldCat pa še niso uspeli uvesti. Za naslednje leto načrtujejo pripra-

vo verzioniranja spletnih storitev CBS in spletnega modula za administracijo uporabnikov in knjižnic ter optimizacijo grupiranja zapisov, spletni vmesnik za prikaz zaloge, storitve *Get OCN service* (direkten dostop iz vmesne baze do baze WorldCat) in optimizacijo SRU API.

POVZETEK SREČANJA

Pri množični integraciji bibliografskih podatkov in vzdrževanju kvalitete gre za revolucijo, evolucijo in paradoks.

Revolucija (pridobivanje virov v realnem okolju; angl. *real world*) vključuje:

- uporabo modelov knjižničnih metapodatkov namesto bibliografskih formatov za izboljšavo poteka dela v knjižnicah,
- splet (kot vzajemno bazo podatkov) z uporabo modela entitet,
- sprejetje novih katalogizacijskih pravil,
- uporabo povezanih podatkov (angl. *linked data*),
- uporabo centralnih virov normativnih metapodatkov (angl. *metadata hubs*),
- uporabo modela grafov,
- migracijo programske opreme na storitve v oblaku.

Evolucija (integracija + sinhronizacija + migracija) vključuje:

- kompleksen sistem upravljanja z metapodatki pri integraciji in sinhronizaciji (kvaliteta podatkov, ujemanje zapisov, združevanje zapisov, grupiranje zapisov, prava podatkov za izmenjavo),
- umeščanje knjižnic kot partnerjev v izobraževanju in znanosti.

Paradoks (zmanjšanje kompleksnosti pri združevanju podatkov) je povezan z:

- enotnimi katalogizacijskimi pravili,
- večjo avtomatizacijo pri katalogizaciji,
- zmanjšanjem pomena katalogizacijskih pravil,
- povečanjem vpliva standardizacije bibliografskih elementov,
- potrebo po sodelovanju, saj samo en partner ne more zagotoviti vseh rešitev (regionalne platforme bodo v uporabi še od 15 do 20 let).

Andreja Krajnc Vobovnik, Gordana Budimir