

ZVUČI KAO *EUROPEANA*

GERDA KOCH

AIT Angewandte Informationstechnik
Forschungsgesellschaft mbH
EuropeanaLocal Austria Servicepoint
DISMARC Servicepoint, LoCloud
Austria
kochg@europeana-local.at
info@dismarc.org

U današnje se vrijeme oko 40% ljudi u dobi između 14 i 30 godina za pretraživanje informacija koristi internetom. Taj postotak stalno raste zahvaljujući sve učestalijoj uporabi pametnih telefona i drugih mobilnih uređaja. U konačnici, internet će biti naš glavni izvor informacija. Stoga je vrlo važno da baštinske ustanove budu vidljive na mreži, a njihovi podaci dostupni i pretraživi. Međutim, nije dovoljno prikazati informacije o kulturnim sadržajima na pojedinačnim mrežnim stranicama. Kako bi se osnažila podrška pronalaženju informacija, važno je također surađivati i umrežavati se, udruživati se putem virtualnih kataloga te se koristiti društvenim medijima. Godine 2006. nekolicina je europskih predsjednika vlade uputila pismo Europskoj komisiji tražeći sredstva za financiranje i pokretanje zajedničke online prezentacije europske kulturne baštine. To je bila polazišna točka za osnutak portala *Europeana*¹ kojim se danas služimo. Prvi prototip *Europeane* zaživio je na mreži potkraj 2008. godine. Otada su

mnogi projekti koje je financirala Europska unija dali svoj tehnološki i sadržajni doprinos portalu. Danas je na portalu dostupno i pretraživo 29 milijuna zapisa iz više od 2200 europskih institucija. *Europeana* je digitalna knjižnica, što znači da portal prihvaća i integrira samo one zapise koji su povezani s pripadajućim digitalnim objektom (tekstom, videom, slikom, zvukom ili 3D objektom).

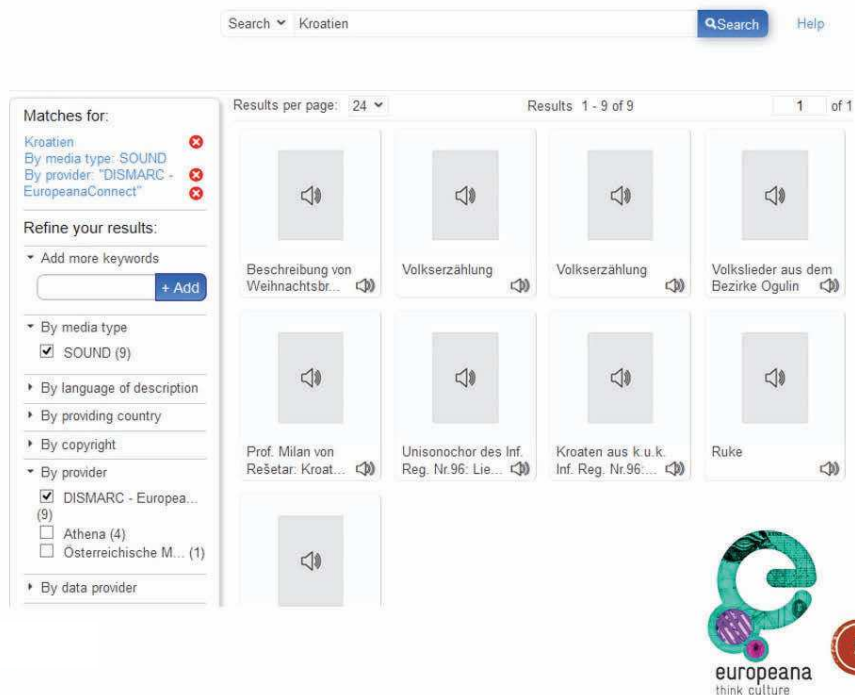
Na sljedećim je slikama predložen način pretraživanja portala. Prvo pretraživanje odnosi se na pojam „Kroatien” koji je unesen u polje za pretraživanje po slobodnom tekstu, a potom je pretraživanje suženo odabirom tipa medija – „zvuk” i davatelja sadržaja – „DISMARC” (sl. 1.). Odabirom identifikacijske sličice (engl. *thumbnail*) u popisu rezultata dobiju se detalji pojedinog zapisa. Na slici 2. vide se svi opisni metapodaci koje je *Europeani* ustupio davatelj sadržaja.

Europeana je tijekom 2013. godine počela obogaćivati svoje metapodatke usklađujući rječnike sa SKOS-om², kako bi se omogućilo povezivanje različitih zapisa. Opisu objekta na slici 2. dodan je i URI (*Uniform Resource Identifier*) za „zvuk” (www.eionet.europa.eu/gemet/concept/7933). Na desnoj su strani prikazane i brze poveznice za fasetirano pretraživanje. Pri dnu lijeve strane prikaza na slici 2. dane su opcije povezivanja s različitim društvenim medijima. Odabir identifikacijske sličice ili naziva davatelja sadržaja (View item at:) vodi do digitalnog objekta ili mrežne stranice baštinske ustanove. Portal *Europeana* pohranjuje samo metapodatke (katalogske podatke) koji su mu

¹ <http://www.europeana.eu> (citirano: 2013-12-13).

² SKOS - Simple Knowledge Organization System, <http://www.w3.org/2004/02/skos/> (citirano: 2013-12-13).

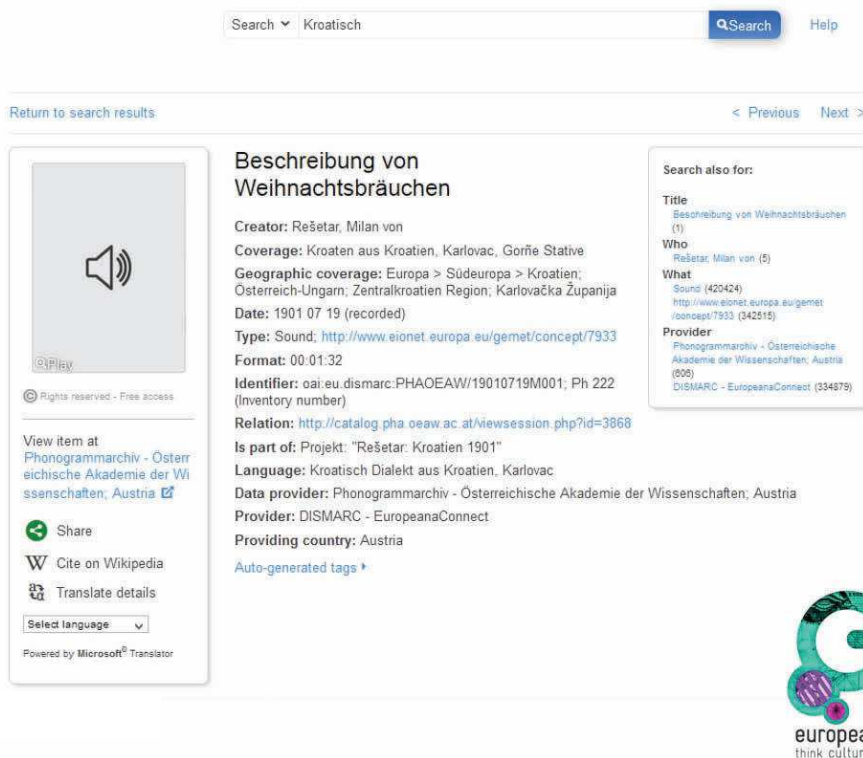
- Filter functionalities



Slika 1.

- Link to digital file

- Social Media




Slika 2.

prosljeđeni. Ti podaci sadržavaju poveznicu na digitalni objekt ili mrežnu stranicu s koje se može pristupiti digitalnom objektu.

Na slici 3. vidi se zapis objekta na mrežnoj stranici institucije – davatelja sadržaja, a može se uočiti i poveznica na digitalni objekt.

Phonogrammarchiv

Österreichische Akademie der Wissenschaften



Suche

Volltextsuche und
Expertensuche

Listenansichten

Instrumente alphabetisch

Instrumente hierarchisch

Instrumente nach Hornbostel-Sachs-Systematik, tabellarisch

Instrumente nach Hornbostel-Sachs-Systematik

Gruppen alphabetisch

Gruppen hierarchisch



Sprachen alphabetisch


Sprachen hierarchisch

Länder alphabetisch

Länder nach Kontinenten

Zurück

Datum (Signatur)	19010719.M001	Audio
Archivnummer(n)	Ph 222	
Inhalt	Beschreibung von Weihnachtsbräuchen	
Dauer (automatisch)	00:01:32	
Projekt	Rešetar: Kroatien 1901	
Kontinent	Europa	
Land	Kroatien	
Land historisch	Österreich-Ungarn	
Region	Zentralkroatien Region , Karlovačka Županija	
Ort	Karlstadt	
Urheber	Rešetar, Milan von	
Gruppen(n)	Kroaten aus Kroatien, Karlovac, Gorње Stative	
Sprache(n)	Kroatisch Dialekt (Kroatisch) aus Kroatien, Karlovac	
Medien/Onlinekatalog		
Archivnummer	Ph 222 (excerpt)	
Titel	Beschreibung von Weihnachtsbräuchen in Serbokroatisch	
Beschreibung	Die vollständige Aufnahme wurde auf OFAW PHA CD 7 : "The First Expeditions 1901 ..." (= Gesamtausgabe der historischen Bestände 1899-1950, Serie 1) publiziert.	
Beschreibung englisch	The complete recording was published in OFAW PHA CD 7 : "The First Expeditions 1901 ..." (= Complete Edition of the Historical Recordings 1899-1950, Series 1).	
Dauer (hh:mm:ss)	00:00:20	
Datei		

Slika 3.

Metapodatkovni format ESE (*Europeanini* semantički elementi, engl. *Europeanina Semantic Elements*), primjenjivan od samih početaka *Europeane*, uglavnom se zasniva na Dublinskome osnovnom skupu metapodataka³ koji je zapravo linearan i prilično ograničen opisni format. Novi standard EDM (*Europeanin* podatkovni model, engl. *Europeanina Data Model*), koji se počeo primjenjivati 2013. godine, omogućuje raznolikije opise predmeta kulturne baštine. Slika 4. daje sažeti prikaz svih preduvjeta koje novi davatelj sadržaja mora zadovoljiti ako se želi priključiti portalu *Europeana*.

³ <http://dublincore.org/documents/2012/06/14/dcmi-terms/> (citirano: 2013-12-13).

Nastavak ovoga rada posvećen je glavnom agregatoru zvučnih zapisa za pristup *Europeani* – portalu *DISMARC*⁴. Taj je portal rezultat projekta što ga je financirala Europska unija radi stvaranja infrastrukture za zbirke zvučnih zapisa i pretraživanje njihovih sadržaja na webu. Projekt je trajao od 2006. do 2008. godine, dakle, završen je prije pokretanja prvog prototipa *Europeane* (kraj 2008. godine).

Za razliku od *Europeane*, *DISMARC* je bio usmjeren na stvaranje interoperabilnih i online pretraživih metapodataka zbirki zvučnih zapisa – njegov primarni cilj nije bilo omogućivanje pristupa

⁴ <http://www.dismarc.eu> (citirano: 2013-12-13).

- Catalog data (Metadata) in electronic format (Public domain)
- Digital sound (text, video, image, 3D) files accessible (CC rights)
- Provide the metadata according to Europeana standard
 - **ESE** (Europeana Semantic Elements) = Dublin Core extended
 - **EDM** (Europeana Data Model) = open, cross-domain Semantic Web-based framework, supports linking of data
 - Main distinctions in EDM:
 - Information on the (real) object (Cultural Heritage Object)
 - Information on the digital representations of the object (Web Resource)
 - Information on both as one logical whole (Aggregation)
 - Contextual Information (Vocabularies)



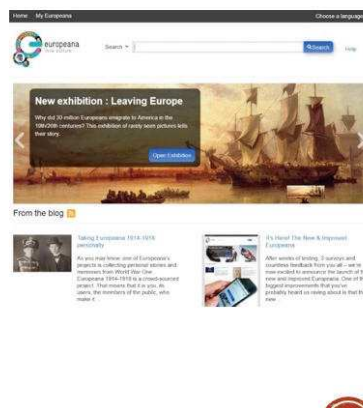
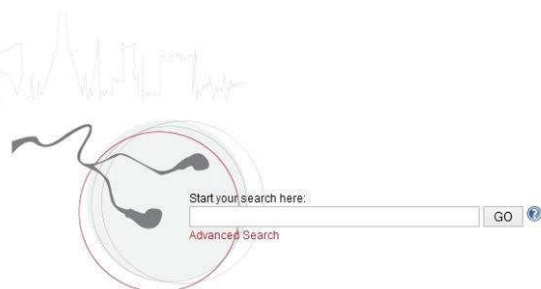
Slika 4.

DISMARC TODAY YEAR 2013

WWW.DISMARC.EU

- 1 Portal (21 Languages)
- 1219 Collections from...
- 1183 Archives/Labels/Museums/Libraries/Private Collectors...
- 2,4 Mio. Metadata records (descriptions)
- 402 306 Metadata records (with digital objects)

- **Audio Provider for Europeana.eu**
(340.000± Records –
Audio, Video, Image, Text)



Slika 5.

SEARCH PORTAL (1)

- Search
 - Single Records (55 mf)
 - Whole collections (39 mf)
- Search options
 - Simple Search
 - Advanced Search
 - Browse
 - Where? > new June 2013
- Index
 - All datafields searchable
- Europeana / DPLA API
 - Search DISMARC and in parallel Europeana.eu
 - Search DISMARC and in parallel the Digital Public Library of America > new April 2013



Slika 6.

samim digitalnim objektima (zvučnim datotekama itd.). Nakon završetka projekta uslijedio je poziv da se *DISMARC* priključi projektu *EuropeanaConnect*⁵, čime je u konačnici postao glavni sustav *Europeana* za agregaciju zvučnih zapisa. (*Europeana* se koristi tzv. agregatorima sadržaja koji pobiru podatke iz pojedinih baštinskih ustanova kako bi se pojednostavnio proces isporuke podataka.) U početku je koordinator projekta *DISMARC* bio Multikulturalni odjel Radija Berlin Brandenburg (RBB), što je omogućilo uspostavu višejezičnog portala koji podržava uporabu 21 jezika. Danas *DISMARC* zajedno administriraju RBB i

AIT (Angewandte Informationstechnik), a portal *Europeana* isporučuje 70% zvučnih sadržaja.

DISMARC je virtualni katalog podataka o zvučnim zapisima koji podržava pretraživanje na razini pojedinačnih zapisa ili na razini cijele zbirke. Pojedinačno je pretraživo svih 55 metapodatkovnih polja za opis na razini zapisa i svih 39 metapodatkovnih polja za opis na razini zbirke zvučnih zapisa.

Alati za pretraživanje i katalog podataka *DISMARC-a* obuhvaćaju 10 višejezičnih rječnika i 15 višejezičnih popisa riječi. Rječnici i popisi riječi dragocjeni su mehanizmi povezivanja heterogenih višejezičnih podataka iz različitih izvora koji omogućuju korisniku jednostavno pregledavanje skupnog kataloga. Tijekom

⁵ <http://www.europeanaconnect.eu/> (citirano: 2013-12-13).

The screenshot displays the DISMARC portal interface. On the left, there is a navigation menu with sections like 'Hauptmenü', 'Suche', 'Kommunikation', and 'Benutzermenü'. The main content area shows search results for the term 'Tanz' in Croatian. A search filter is set to 'Croatian'. The results list includes 'Tanz' and 'Tanzprogramm'. A 'Wordlists' section on the right shows a list of instruments: harmonika, duduk, alto truba, alto saksofon, balafon, banjo, barton saksofon, bas, and bas klarineta. A text box in the center of the screenshot states: '10 Vocabularies 16 word lists'.

Slika 7.

preuzimanja podataka na *DISMARC-ovu* platformu pojmovi iz višejezičnog rječnika automatski se dodaju metapodacima zapisa. Pojedini partneri u projektu, tj. davatelji sadržaja, također se koriste rječnicima za potrebe lokalne katalogizacije. Popisi riječi i rječnici mogu se koristiti pri pregledavanju baze podataka, ali su ujedno integrirani s alatima višejezičnog pretraživanja.

Ako korisnik aktivira funkciju višejezičnog pretraživanja, moguće je, primjerice, upisom njemačke riječi „Tanz” pronaći i zapise koji sadržavaju riječ „danza”. Korisnik bira jezik pretraživanja!

Godine 2013. *DISMARC-u* su prirodane dvije nove funkcije pretraživanja: alat za pregledavanje prema geografskom položaju i sučelje (Application Programming Interface, API) prema DPLA – Digital Public Library of America⁶. Integraci-

jom sučelja prema DPLA (DPLA API) omogućeno je istodobno pretraživanje *DISMARC-a* te digitalnih knjižnica *Europeane* i DPLA-e. Popisi rezultata pretraživanja iz *Europeane* i DPLA-e prikazuju se na posebnim karticama (engl. *tabs*), uz popis rezultata iz *DISMARC-a*. Upotrebom aplikacije *Google Maps* omogućeno je pregledavanje prema geografskom položaju. Na karti se mogu pronaći zapisi i zbirke iz pojedinih zemalja.

Ujesen 2013. godine započela je dogradnja portala suradničkom platformom i taj je posao još uvijek u tijeku. Ideja o kiranju suradničke platforme rezultat je učestalih korisničkih zahtjeva i poruka primljenih elektroničkom poštom tijekom proteklih godina. Alati nove suradničke platforme omogućit će jednostavan unos podataka o zbirkama zvučnih zapisa te njihovu isporuku *DISMARC-u* i *Europeani*. Alati će biti posebno prilagođe-

⁶ <http://dp.la/> (citirano: 2013-12-13).

ni malim institucijama i privatnim kolekcionarima. Bit će ponuđen izbor funkcija za preuzimanje podataka i omogućeno korištenje *DISMARC-ovim* višejezičnim rječnicima za potrebe katalogiziranja. Predviđeno je da će se suradnička platforma pokrenuti tijekom 2014. godine.

Naposljetku, željela bih zaključiti kratkim prikazom jednoga od najnovijih *Europeaninih* projekata – projekta *LoCloud*⁷. Projekt nije osmišljen samo za zbirke zvučnih zapisa, već općenito za male i srednje baštinske ustanove. Iniciran je u ožujku 2013. godine radi tri glavna cilja:

- podrške baštinskim ustanovama u predstavljanju njihovih sadržaja na *Europeani*
- korištenja računarstva u oblaku (engl. *cloud computing*) kako bi se osigurale daljinske usluge i alati kojima bi se smanjile tehničke, semantičke i stručne prepreke
- omogućivanja jednostavnije agregacije digitalnih sadržaja iz malih i srednjih baštinskih ustanova.

Osim *LoClouda*, razvija se i tzv. sestrinski projekt – *EuropeanaCloud*, koji izravno koordinira *Europeana* zaklada. Za potrebe *Europeane* i njezinih agregatora projektom *EuropeanaCloud* uspostaviti će se sustav utemeljen na računarstvu u oblaku.

Općenito, usluge računarstva u oblaku (engl. *cloud services*)⁸ daljinske su usluge dostupne s udaljenih lokacija, kojima se putem interneta mogu koristiti lokalni klijenti. Razlikujemo nekoliko modela usluga računarstva u oblaku:

- softver kao usluga (engl. *software as a service*, SaaS): *Dropbox*, *Google Apps*, *Evernote*...

- platforma kao usluga (engl. *platform as a service*, PaaS): računalna platforma za razvoj i korištenje softverskih rješenja, uključujući i operacijski sustav

- infrastruktura kao usluga (engl. *infrastructure as a service*, IaaS): računala – fizička i (sve češće) virtualna.

Ako su daljinske usluge javno ponuđene, riječ je o javnom oblaku (engl. *public cloud*). Ako su, pak, te usluge namijenjene pojedinoj organizaciji ili tvrtki, koje ona može i sama kontrolirati u vlastitome računalnom centru, govori se o privatnom oblaku (engl. *private cloud*).

U sklopu projekta *LoCloud* razvit će se model softvera kao usluge⁹, čije će aplikacije biti prilagođene potrebama malih i srednjih baštinskih ustanova. Planiran je razvoj ovih alata:

- naprednih geolokacijskih alata
 - alata za obradu prirodnog jezika (engl. *natural language processing*, NLP) radi obogaćivanja metapodataka (engleski/španjolski)
 - suradničkih alata za razvoj višejezičnih rječnika
 - suradničkih alata za razvoj geografskog leksikona (indeksa) povijesnih naziva mjesta
 - alata za korištenje Wikimedije u projektima s velikim brojem vanjskih suradnika (engl. *crowdsourcing projects*).
- Željela bih zaključiti pozivom da se uključite u online mreže i stručne zajednice. Novi su partneri uvijek dobrodošli u mreže predstavljene u ovome radu, uz podršku agregatora koji omogućuju:

⁷ <http://www.locloud.eu/> (citirano: 2013-12-13).

⁸ http://en.wikipedia.org/wiki/Cloud_computing (citirano: 2013-12-13).

⁹ http://en.wikipedia.org/wiki/Software_as_a_service (citirano: 2013-12-13).

- konverziju metapodataka (lokalnih metapodataka u EDM)
- obogaćivanje metapodataka
- pohranu metapodataka (uz pohranu zvučnih zapisa u *DISMARC-u*)
- ustupanje metapodataka *Europeani*.

Ako vas zanimaju neki od alata *LoClouda*, pridružite se probnoj fazi projekta tijekom 2014. godine.

Sa engleskog jezika prevela Maja Šojat Bikić

SOUNDS LIKE EUROPEANA

Since its first online launch in late 2008 Europeana has become one of the biggest cross-domain and cross-country digital libraries for cultural heritage objects in the world. To date the Europeana network includes more than 500 members and the approx. 28 Mio records accessible now via the Europeana portal at www.europeana.eu stem from around 2,200 contentcontributing partners across Europe. In its business plan for 2013 Europeana states that in particular the audio and audiovisual content gets most attention from the end-users. This is an important fact for the promotion and visibility of music archives.

As audio-visual content makes up to date only about 3% of the Europeana database it is one of the goals for year 2013 and the future to focus on the ingestion of audio and audio-visual data. In 2008 the DISMARC portal, www.dismarc.eu, launched as well. This portal was developed during a two years EU-funded project and is today jointly administered by Rundfunk Berlin-Brandenburg (RBB, Berlin) and Angewandte Informationstechnik Forschungsgesellschaft mbH (AIT, Graz). It was the first international online platform to provide access to over 2 Million audio records contributed by music archives, music libraries, phonomuseums, and music experts all over Europe and even across the Atlantic. Since 2009 DISMARC serves as the main audio aggregator to Europeana, 70% of the sound records in Europeana are delivered by DISMARC. Its content and user base is constantly growing and several new features have been added to the platform in spring 2013 (eg. the API of the Digital Public Library of America, a geographical map based browsing facility). The presentation will introduce to the ingest workflow for audio data to Europeana via the DISMARC platform and will describe the new collaborative functionalities and cloud technologies that shall be implemented in future.