

Realizacija audio-biblioteke Univerziteta u Novom Sadu

Dragiša M. Mišković, Milan J. Gnjatović, Nikša M. Jakovljević, Vlado D. Delić

Sadržaj — U radu su opisani izabrani aspekti specifikacije i dizajna centralne audio-biblioteke Univerziteta u Novom Sadu (CABUNS). Cilj projekta je unapređenje obrazovnog procesa kreiranjem audio-izdanja knjiga i predavanja. Aplikacija je realizovana kao veb-portal koji primenjuje automatsku sintezu govora za srpski jezik i generiše adekvatne audio-vizuelne sadržaje. U radu su istaknute osnovne karakteristike primenjene arhitekture, i mogućnosti za dalja unapređenja sistema.

Gljučne reči — audio-biblioteka, veb-server, sinteza govora, ASP.NET, MVC.

I. UVOD

RAZVOJ interneta u pravcu globalne mreže informacija uvodi brojne novine u svakodnevni život. Već duže vremena, tradicionalne biblioteke se uključuju u proces digitalizacije koja za cilj ima smeštaj bibliotekskog fonda na nove medijume kao što su magnetni i optički diskovi. U novije vreme, razvoj mrežnih tehnologija generisao je još jedan stepen slobode koji u potpunosti izbacuje potrebu za fizičkim medijumom na koji bi se smeštali preuzeti bibliotekski sadržaji. U pitanju je virtuelizacija podataka omogućena pristupom preko Interneta. Prednosti ovakve virtuelizacije su brojne:

- bibliotekski materijal je dostupan korisnicima širom sveta,
- u svakom momentu, elektronskom dokumentu može pristupiti više korisnika istovremeno,
- ukidaju se ograničenja nametnuta radnim vremenom biblioteke,
- olakšava se proces pretrage materijala, itd.

Kao rezultat, veliki broj nacionalnih biblioteka ulaže značajne napore u kreiranje elektronskih digitalnih zbirki koje bi, preko Interneta, bile dostupne pojedinačnim korisnicima. Po obimu digitalizovanog materijala se ističu kongresna biblioteka u Vašingtonu, koja je, po podacima iz 2002. godine [1], imala oko tri miliona digitalnih dokumenata, Britanska nacionalna biblioteka u Londonu i

Francuska nacionalna biblioteka u Parizu (takođe sa impresivnim zbirka digitalnih dokumenata).

U Srbiji je najveći iskorak u ovom smeru napravljen u okviru projekta „Digitalna Narodna biblioteka Srbije“, koji se realizuje od 2003. godine. Ova digitalna biblioteka trenutno sadrži više od milion digitalnih dokumenata, a projekat je 2014. godine ušao u fazu ozvučavanja digitalnih objekata. Na ovaj način, započeo je novi proces transformacije biblioteka, pri čemu se komunikacija sa korisnicima dodatno obogaćuje - do sada dominantna vizuelna priroda ove komunikacije se upotpunjuje govornom komponentom. Kao rezultat, svi koji više vole da slušaju nego da sami čitaju, a posebno osobe sa određenim invaliditetom dobijaju mogućnost da pristupe audio-sadržajima biblioteka, ali je potrebno istaći i druge prednosti ovakvog pritupa:

- u kombinaciji sa mobilnim uređajima, moguće je preslušavanje delova knjiga u situacijama dok se obavljaju neke druge aktivnosti,
- moguće je kombinovanje audio i vizuelnog pristupa adekvatnim sadržajima.

Ovi aspekti su bili motiv za razvoj centralne audio-biblioteke Univerziteta u Novom Sadu (CABUNS). Cilj projekta je unapređenje obrazovanja primenom inovativnih resursa zasnovanih na automatskoj sintezi govora na osnovu teksta i savremenim informaciono-komunikacionim tehnologijama [2]. CABUNS je specifikiran kao centralizovani tehnički sistem koji će omogućiti studentima pristup audio-izdanjima knjiga i predavanja, u formatu prikladnom za različite uređaje.

II. ARHITEKTURA SISTEMA

Jedan od funkcionalnih zahteva prilikom projektovanja audio-biblioteke je potpuna dostupnost servisa na različitim platformama (mobilni telefoni, tablet-uređaji, računari), sa različitim operativnim sistemima. Usled toga, izabran je dizajn sistema zasnovan na veb-interfejsu.

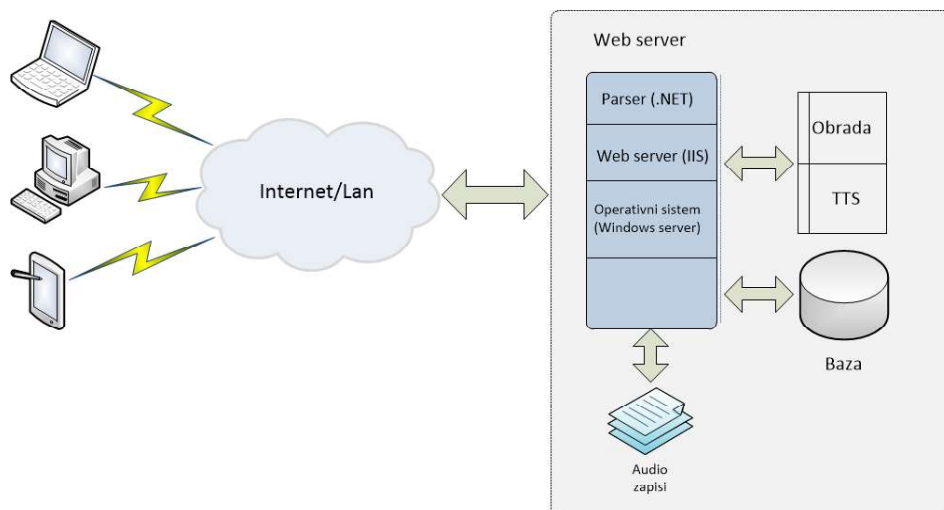
Slika 1 prikazuje arhitekturu sistema kome korisnici pristupaju preko dinamičke veb-strane. Za upravljanje korisničkim zahtevima nadležan je IIS veb-server (engl. *Internet Information Service - IIS*), koji se zasniva na skupu standardnih formata i protokola. HTML (engl. *Hypertext Markup Language*) omogućava formatiranje veb-strana, dok se HTTP (engl. *Hypertext Transfer Protocol*) primenjuje kao osnovni komunikacioni protokol za razmenu informacija između klijenata i veb-servera.

D. M. Mišković, Fakultet tehničkih Nauka, Novi Sad, Srbija (e-mail: dragisa@uns.ac.rs).

M. J. Gnjatović, Fakultet tehničkih Nauka, Novi Sad, Srbija (e-mail: milangnjatovic@uns.ac.rs).

N. M. Jakovljević, Fakultet tehničkih Nauka, Novi Sad, Srbija; (e-mail: jakovnik@uns.ac.rs)

V. D. Delić, Fakultet tehničkih Nauka, Novi Sad, Srbija; (e-mail: vdelic@uns.ac.rs)

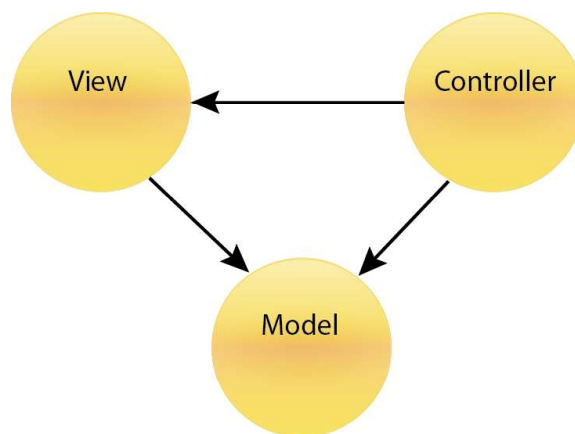


Slika 1. Arhitektura audio-biblioteke. Korisnički interfejs je realizovan kao dinamička veb-strana.

Serverska strana ovog interfejsa realizovana je ASP.NET tehnologijom (engl. *Active Server Pages - ASP*) [3]. Tokom korisničkih poseta portalu audio-biblioteke, HTTP-zahtevi se prosleđuju IIS-serveru, koji ih identifikuje i usmerava do ASP-sloja, na kome se vrše obrada, pronalaženje podataka u bazi i slanje adekvatno formatiranog odgovora klijentu, takođe posredstvom IIS-servera. Za realizaciju ovog protokola je primenjen tzv. MVC-pristup (engl. *Model-View-Controller*). Ključna karakteristika ovog pristupa je da, u odnosu na druge arhitekture, obezbeđuje viši nivo apstrakcije, razdvajanjem funkcionalnosti aplikacije na tri ključne komponente (Sl. 2):

- **Model** – zadužen za upravljanje domenom podataka. Ovaj deo realizuje komunikaciju sa bazom i definiše pravila kojima se uređuje odnos između objekata definisanih u aplikaciji.
- **Prezentacija** – obezbeđuje adekvatni prikaz podataka u okviru korisničkog interfejsa (engl. *user interface - UI*).
- **Kontroler** – vrši obradu korisničkog zahteva i predstavlja vezu između korisničkog interfejsa i modela aplikacije. Nakon obrade zahteva i pribavljanja informacija od modela, kontroler selektuje odgovarajuću formu prezentacije traženih podataka.

Na ovaj način, ASP-MVC-pristup omogućava nezavisni razvoj različitih aspekata aplikacije i definiše način za uspostavljanje veza između njih. Sa stanovišta primene u okviru projekta CABUNS, ovo je veoma značajno, posebno ako se ima u vidu zahtev za funkcionalnošću sistema na različitim uređajima. Opisani mehanizam apstrakcije omogućava da se eventualna prilagođenja za različite platforme vrše na nivou prezentacione logike, dok ostali delovi sistema ostaju nepromenjeni.



Slika 2. Komponente MVC-pristupa i veze između njih.

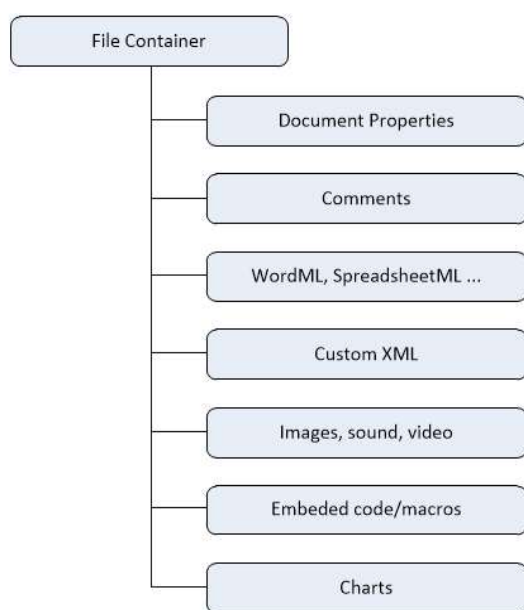
III. OBRADA ZAHTEVA

A. Obrada dokumentata

Sistem trenutno prihvata dve vrste datoteka iz paketa Majkrosoft Ofis (engl. *Microsoft Office*): prezentacije (*PowerPoint*) i dokumenata (*Word*). Mogućnost prosleđivanja datoteka sistemu je ograničena na registrovane korisnike (tj. profesore Univerziteta u Novom Sadu).

Potrebno je naglasiti da obrada datoteka može biti vremenski zahtevna (u zavisnosti od njihove veličine), pa se celokupna procedura obrade obavlja kao pozadinski proces. U okviru ovog procesa se vrše (i) parsiranje dokumenta kako bi se izdvojio tekst koji će se slati na sintezu govora i (ii) izdvajanje sadržaja značajnih za vizuelni prikaz u okviru veb-strane. Kod prezentacija (*PowerPoint*), audio-zapis se generiše na osnovu sadržaja beleški dodatih svakom slajdu. Sistem takođe zadržava vizuelni prikaz slajda, i prikazuje ga u okviru veb-strane koja odgovara datoj lekciji. Kod dokumenata (*Word*), obrada je nešto drugačija – celokupni tekst iz dokumenta se izdvaja, i na osnovu njega se sintetiše govor, a vizuelni sadržaj se formira na osnovu slika i tabela sadržanih u dokumentu.

Obrada dokumenata se zasniva na *Open Office XML* formatu koji je 2007. godine usvojen kao standard za *Microsoft Office* dokumente. Za razliku od ranijeg, binarnog formata za predstavljanje dokumenata, ovaj format interno kreira modularnu strukturu dokumenta prikazanu na slici 3. Komprimovana XML-forma objedinjuje različite delove dokumenta koji su povezani međusobnim relacijama. Kao posledica, moguće je pristupiti pojedinim delovima dokumenta (slike, tabele, beleške kod prezentacija, itd.) bez korišćenja *Office* programa. Parsiranje sadržaja i kreiranje odgovarajućeg audio-vizuelnog prikaza realizuje se uz primenu *Open XML* razvojnog alata (engl. *Office Open XML Software Development Kit – Open XML SDK*).



Slika 3. Interna struktura *Microsoft Office* dokumenata. Komprimovana XML forma objedinjuje različite delove dokumenata i njihove međusobne relacije.

Proces obrade dokumenata i generisanja zvučnog zapisa je odvojen od funkcionalnosti servera, i realizovan je u vidu zasebnog modula. Nakon pristizanja novog dokumenta, server inicira njegovu obradu unutar ovog modula, a kao rezultat dobija listu fajlova koje smešta u bazu podataka i povezuje sa odgovarajućim predavanjem. Izdvojeni proces obrade dokumenata omogućava veću fleksibilnost u procesu razvoja aplikacije (npr. izmene u sintetizatoru govora ne zahtevaju promene unutar servera).

B. Generisanje audio-zapisa

Početna verzija CABUNS-portala koristi tzv. konkatenativnu sintezu razvijenu u okviru istraživačkog tima AlfaNum prethodnih godina. Ovo je prvi sintetizator govora na srpskom govornom području, a njegovo funkcionisanje se zasniva na korišćenju velike govorne baze iz koje se, nakon analize teksta, selektuju i spajaju različiti govorni segmenti [5][6]. Priroda ovakvog tipa sintetizovanja govora zahteva postojanje velikog rečnika i razvoj tehnika sintaksne analize za svaki jezik. Do sada su razvijeni sintetizatori sa dva govornika za srpski i jednim govornikom za hrvatski jezik.

Uzimajući u obzir ciljeve projekta CABUNS, kao jedan od zahteva se nameće fleksibilnost sintetizatora govora po pitanjima promene glasa, karakteristika govornika, govornog stila, itd. Zbog toga su značajna istraživanja u oblasti tzv. parametarske sinteze govora. Osnovna ideja ovih sintetizatora se zasniva na kreiranju kontekstno zavisnih modela fonema, koji se kasnije koriste pri generisanju govornog signala [7]. Postoje dva pristupa kreiranju ovih modela: primenom skrivenih Markovljevih modela (ovo je pristup zastupljen kod prvih parametarskih sinteza) i primenom neuronskih mreža. Drugim pristupom se postižu bolji rezultati, pošto neuronske mreže mogu da modeluju složenije kontekstne zavisnosti i predstavljaju kompaktniji alat za eksploataciju.

U okviru istraživačko-razvojnog centra AlfaNum je u završnoj fazi razvoj sintetizatora govora za srpski jezik zasnovan na neuronskim mrežama. U toku je eksperimentalna primena ovakve sinteze za potrebe CABUNS projekta. Kao rezultat, očekuje se da audio-sadržaji sa predavanja sadrže govor koji je prijatniji za slušanje, sa promenljivom intonacijom u cilju naglašavanja pojedinih delova, sa različitim bojama glasa i drugim karakteristikama prirodnog govora.

IV. ZAKLJUČAK

U radu je prikazan deo istraživačko-razvojnih aktivnosti u okviru projekta „Centralna audio-biblioteka Univerziteta u Novom Sadu“. Prikazani su ključni elementi razvijenog portala audio-biblioteke, uz osvrt na izabrane tehnologije i arhitekturu celokupnog sistema. Naglašena je modularnost kao osnovna karakteristika izabranog pristupa u razvoju, koja treba da omogući fleksibilna poboljšanja servisa.

Budući rad na projektu uključuje sledeća unapređenja sistema:

- primena parametarske sinteze govora prilikom generisanja audio-sadržaja,
- proširenje primene sistema za više jezika,
- podrška za različite format audio-vizuelnog prikaza predavanja (npr. video-sadržaji),
- prilagođenja za mobilne platforme,
- omogućavanje interakcije između studenata i profesora preko portala audio-biblioteke.

ZAHVALNICA

Istraživanje je realizovano u okviru projekta „Centralna audio-biblioteka Univerziteta u Novom Sadu“, finansiranog od strane Pokrajinskog sekretarijata za visoko obrazovanje i naučnoistraživačku delatnost AP Vojvodine, projekta „Razvoj dijaloških sistema za srpski i druge južnoslovenske jezike“ (TR32035), finansiranog od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije, kao i EUREKA projekta DANSPLAT (E19944).

LITERATURA

- [1] V. Injac, „Digitalne biblioteke u svetu“, Pregled nacionalnog centra za digitalizaciju, pp. 26-39, 2002.
- [2] V. Delić, M. Sečujski, D. Mišković, D. Nikolić, „Koncept centralne audio-biblioteke Univerziteta u Novom Sadu“, *XXII skup Trendovi*

razvoja, TREND, Zlatibor, ISBN 978-86-7392-795-9, pp. 45-48, 2016.

- [3] Overview of ASP.NET Core MVC, Microsoft, <https://www.asp.net/mvc>.
- [4] Office Open XML, Microsoft, <http://officeopenxml.com/>.
- [5] M. Sečujski, R. Obradović, D. Pekar, Lj. Jovanov, V. Delić, "Alfanum System for Speech Synthesis in Serbian Language", In *Proc. of 5th Conf. Text, Speech & Dialogue*, Brno, Czech Republic, 2002.
- [6] V. Delić, M. Sečujski, D. Pekar, N. Jakovljević, D. Mišković, "A Review of Alfanum Speech Technologies for Serbian, Croatian and Macedonian", *Proc. Int. Language Technologies Conference IS-LTC'06*, Ljubljana, Slovenia, pp. 257-260, 2006.
- [7] T. Delić, M. Sečujski, S. Suzić, "A review of Serbian parametric speech synthesis based on deep neural networks", *Telfor Journal, Telecommunications Society and Academic Mind*, Belgrade, ISSN: 1821-3251, Vol. 9, No. 1, pp. 32-37, 2017.
- [8] H. Zen, A. Senior, M. Schuster, "Statistical parametric speech synthesis using deep neural networks", *ICASSP*, pp. 7962-7966, 2013.

ABSTRACT

The paper reports on the selected aspects of the specification and design of the Central Audio Library of the University of Novi Sad (CABUNS). This project aims at improving the educational process by generating audio books and lectures. The underlying application is designed as a web-portal that incorporates the functionalities of speech-to-text synthesis for the Serbian language and generation of the accompanying visual contents. The paper discusses the system's architecture, and possibilities for further improvement.

DEVELOPMENT OF THE AUDIO LIBRARY OF THE UNIVERSITY OF NOVI SAD

D. Mišković, M. Gnjatović, N. Jakovljević and V. Delić