Преглед НЦД 13 (2008), 41–48

Dragoljub Pokrajac,
(Delaware State University,
Dover, DE USA)
Milan P. Milovanović, and **Vesna Aleksandrović**
(National Library of Serbia)

TOWARDS SERBIAN NATIONAL SOUND ARCHIVES—INFORMATION SYSTEM FOR STORAGE, ACCESS AND RETRIEVAL OF RESTORED SOUND RECORDINGS

Abstract: The paper provides overview of a database for storage, access and retrieval of sound recordings for an audio archive. We concentrate on functionality of the system, and provide discussion of database schema and web-access using php.
Key words: databases, php, postgresql, sound recordings,

1. Introduction

Digitized and restored vintage sound material should be preserved and made available for users. In this study, we provide an overview of the proposed information system for storage, search and retrieval of audio material in future Serbian National Sound Archives. We concentrate on functionality of the system, but we do not discuss pertinent issues such as physical protection of the digital content (back-up copies, uninterrupted power supplies, physical security, etc), performance (the number of concurrent users that can be supported), etc. However, such issues are important when designing any real-world information system and must be taken into consideration in this particular case, too.

2. Design of information system

When designing the information system, we are governed with the required functionality. Hence, the system should make possible search by a selected criteria: performer, type and title of performance, record company, etc. The system is to provide storage and presentation of restored sound material, as well as meta-data (data about the recorded sound!), visual information (e.g., digitalized pictures of record labels or sleeves), links to external web-sites, etc. The principal users of our system are researchers, library patrons, casual users, radio-stations, as well as general public (through external web interface). Major design requirements for the system includes robustness, scalability (the system should be designed to easily support extensions of the records' collection), flexibility (should be capable of storing information related to different kinds of records, recording media, etc.)

The proposed information system consists of the web server and the database server, both to be situated at National Library of Serbia. The user sends a request to the information system from a standardized web-based graphical user interface on a client computer. The client computer should be connected to the Internet, so that the user's request can be transmitted to the web-server. The client's computer communicate user's request through html and java script. The web server, running under Linux open source environment, sends request to the

database server, which executes posgreSQL database management system (DBMS). These requests are sent through SQL queries synthesized in php script. The results of queries are transmitted back from the database server and submitted to a php script, which prepares html document with presentation of the results suitable for the users' needs. The user can visualize textual description of retrieved material, listen to retrieved sound and visualize graphical material such as pictures of records' labels, etc (see Fig. 1).

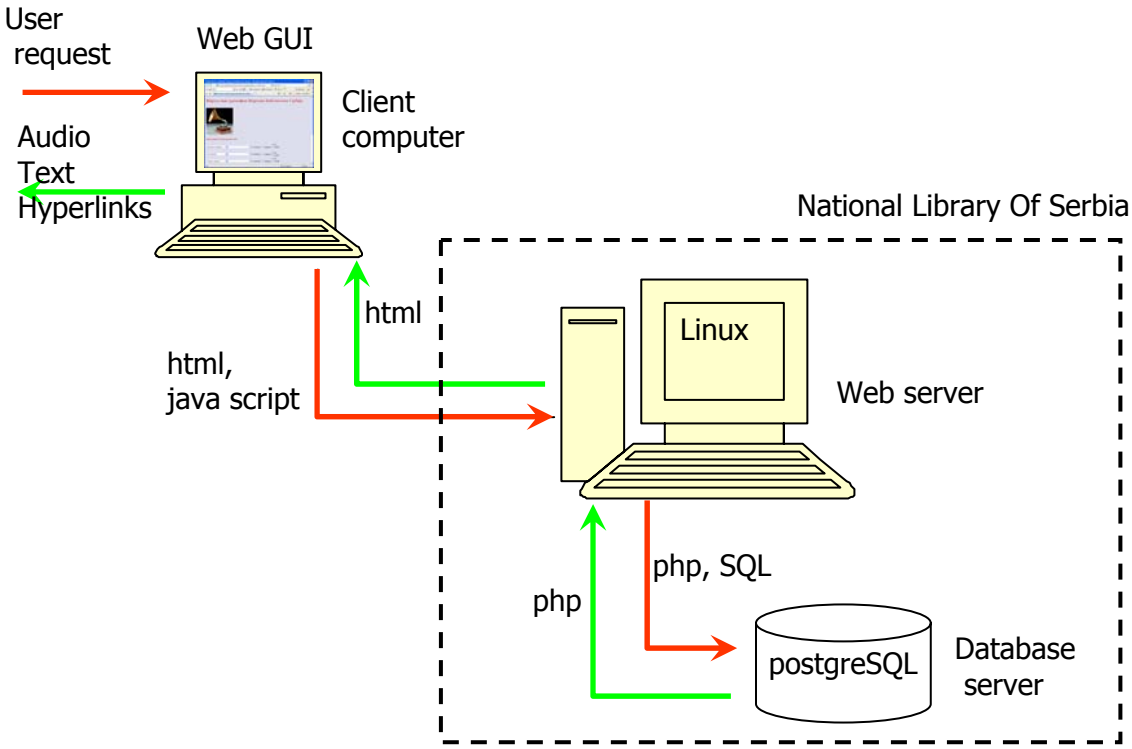


Fig. 1. Design paradigm of the information system.

To support efficient storage and retrieval of data in the information system, we designed an object-relational database that contains textual, graphical and sound information related to recordings, records, labels, performers, etc. The extended entity-relational diagram (Elmasri, Navathe, 2003) of the database is shown at Fig. 3. As we can see, the major part if "record side" (corresponding to a recording from one side of the record), which is related to "record", "orchestra" and "performer". The "record" entity is itself related to "manufacturer" and "library item", where a "library item" can be either "single item" (corresponding to a single record) or "collection" (corresponding to a record album with a single signature number). A "record side" is characterized by matrix number, condition, type of performance, title, multivalued attribute language (a song can be performed in two languages!), data about recording session, data about digitalization, etc. Records are characterized by record size, type, speed, number, condition, medium, etc. A recording side can have multiple performers. Also, a side can have multiple authors, which can have different roles in authoring the specific musical piece. For authors and performers, we keep biographical information and link to an external web site (if exists).

The proposed information system interfaces with end-users by means of graphical user interface implemented in standard html language, and accessible through an html-compliant web-browser, see Fig. 2. The interface is fully implemented in Cyrillic alphabet (using the Unicode standard) and offers search capabilities via up to four criteria selected from the following: Performer's last name, performer's first name, language, manufacturer, label, year of

pressing, title, conductor's last name, composer's last name, orchestra name, type of performance. The search criteria can be combined based into more complex predicates using "and" and "or" operations. Each search predicate can be exact (e.g., "last name is Мијатовић") or approximate (e.g., "orchestra name contains Philadelphia"). A user may choose display of the information related to: records, record sides and performers. In addition, an advanced user may directly use SQL language to submit arbitrary query to the database.

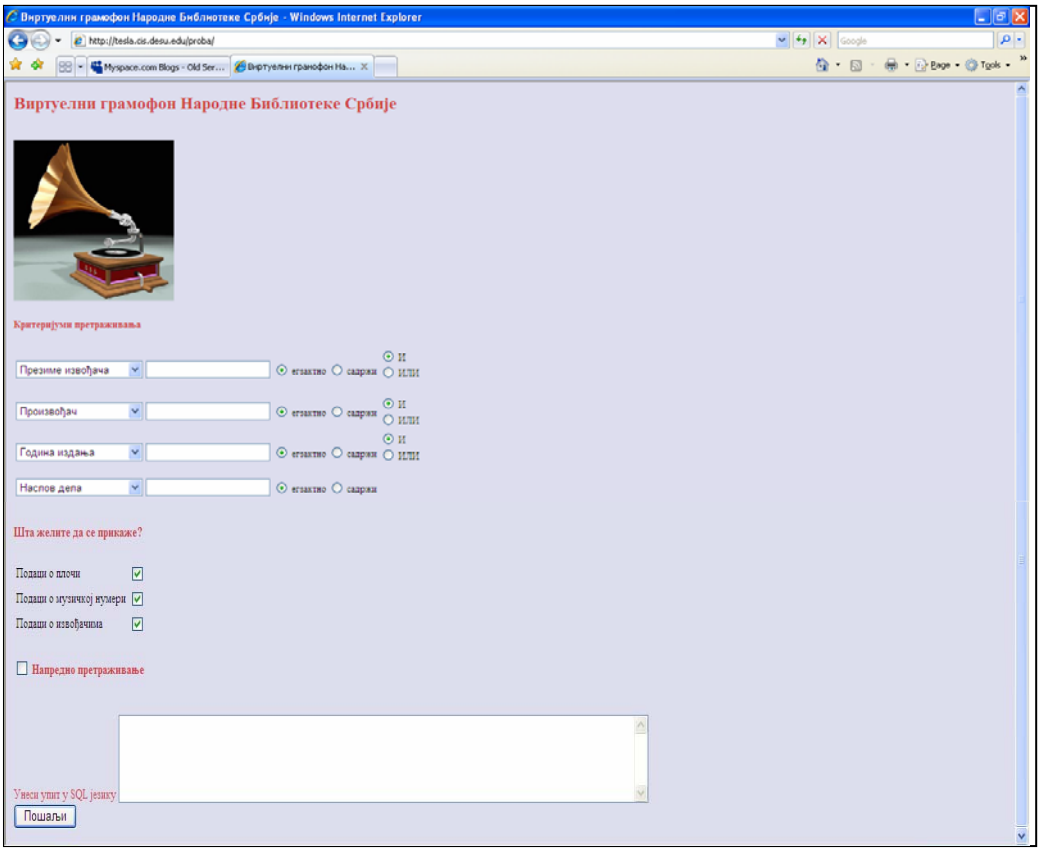


Fig. 2. Front-end: web graphical user interface.

The database has been implemented in postgreSQL, a modern open-source DBMS that can provide compatibility with the recent standards of SQL language (for data definition and manipulation) and support transaction processing (when multiple users try to access the data at the same time) and multiple languages/alphabets through a unicode standard (Geschwinde, Schönig, 2002). A major challenge, to implement disjunctive total specialization of "library items" into "single items" and "collection items" is resolved using work proposed by Pokrajac et al (2004).

SQL queries are created dynamically in php script, based on fields and field values from the web-page. A standard php connection and retrieval of results into an html table is applied (Fig. 4)

Utilization of the information system. As an example of possible utilization of the information system, let us consider a user that is interested in retrieving all recordings belonging to a specific recording type, e.g., folk music. As a result of this query, shown at Fig. 5, the information system will retrieve information in tabular form, as shown in Fig. 6. As shown, for each retrieved recording side, title, matrix number and other relevant information are provided. The user may listen to a recording by clicking at the "gramophone" icon in the corresponding row.

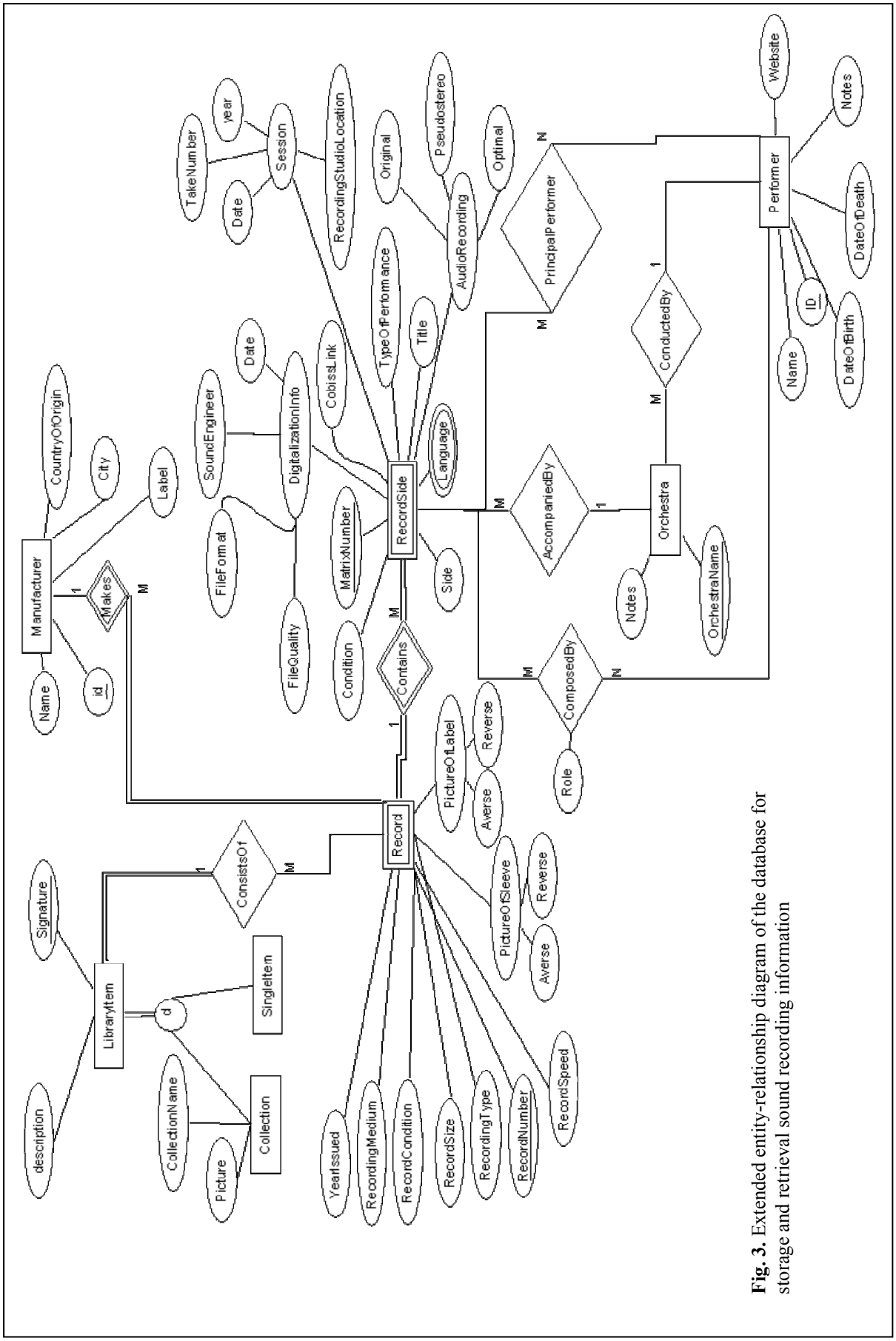


Fig. 3. Extended entity-relationship diagram of the database for storage and retrieval sound recording information

```
$connstr="dbname=library user=pokie";
$dbh=pg_connect($connstr);
$SQLquery=$_POST["SQLquery"];
$stat=pg_query($dbh,$SQLquery);
$rows=pg_numrows($stat);
$columns=pg_num_fields($stat);
 $table="<table border=3><tr>\n"; //initializes table
 for ($j=0;$j<$columns;$j++)
 {
 $ime=pg_fieldname($stat,$j); //retrieves name of j-th returned field
//(very important function!)
 $table.="<th>".$ime."</th>";
 }
```

Fig. 4. A core of php script to submit query to a database and retrieve results

As an example of a more complex query, we demonstrate retrieval of sound recordings and corresponding records composed by Brahms, conducted by Stokowski and performed by orchestra containing word "Philadelphia" in its name, Fig. 7. As a result, two recordings stored in the database are displayed, along with their labels and sound files, Fig. 8.

Finally, we demonstrate an "or" query that retrieves records and performers' data belonging to two different performers, See. Fig. 9 and Fig. 10.

Conclusions

Currently, a test version of the information system is implemented that provides web-interface (available at tesla.cis.desu.edu/proba). Php and html code is implemented and tested and the database is populated with test data from actual holdings of National Library of Serbia. Future work includes: improving GUI using Web 2.0 principles (Kroski, 2008); populating the database with data about all the digitalized records; acquisition of database server at National Library of Serbia to house data and launching and testing the proposed system in real-world environment.

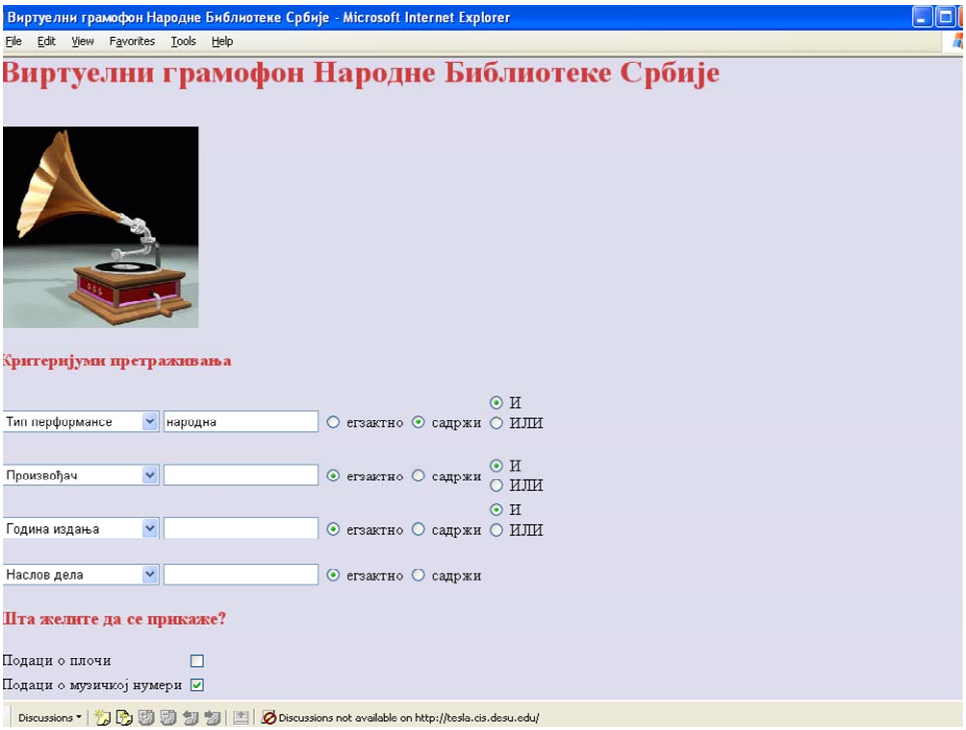


Fig.. 5. Querying of recording sides belonging to folk music.

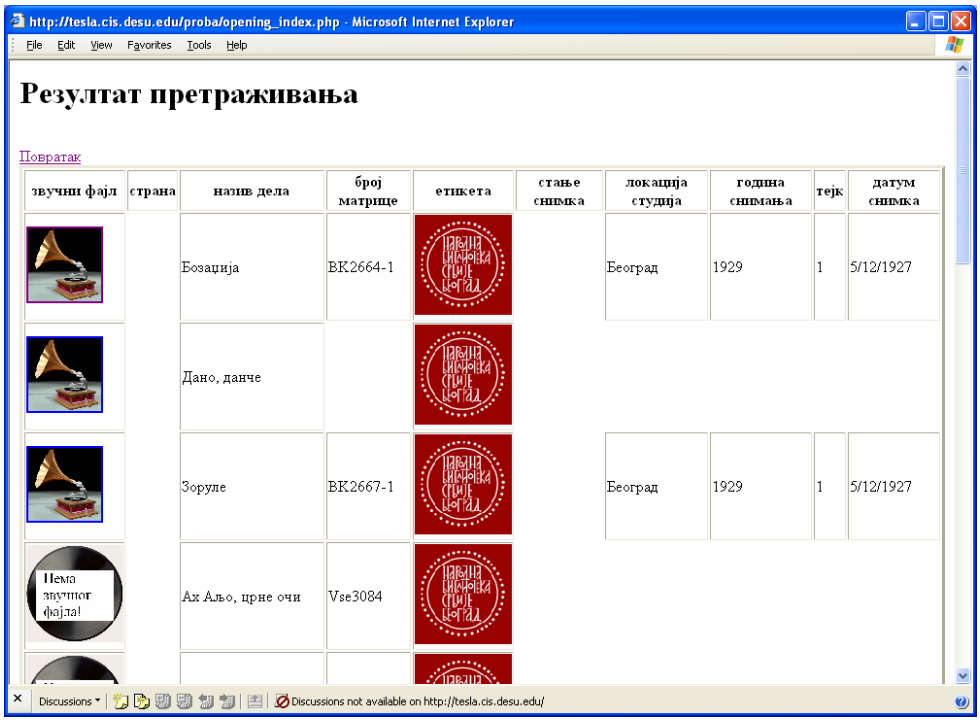


Fig. 6. Tabular result of search from Fig. 5.

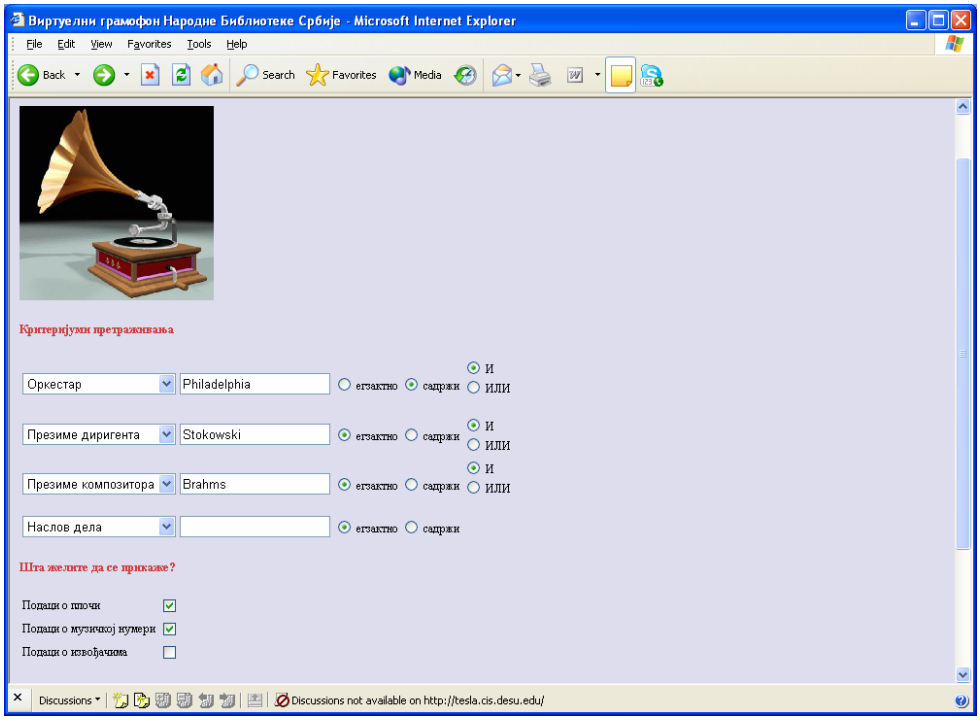


Fig. 7. Complex retrieval combining three predicates.

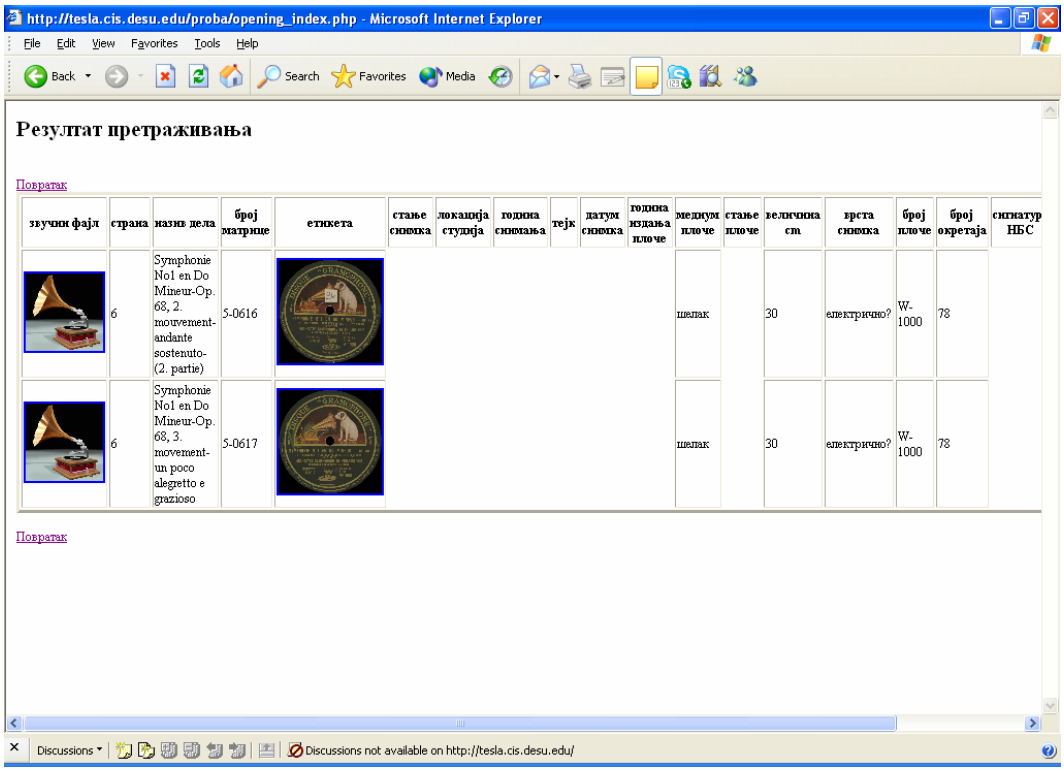


Fig. 8. Results of complex retrieval from Fig. 7.

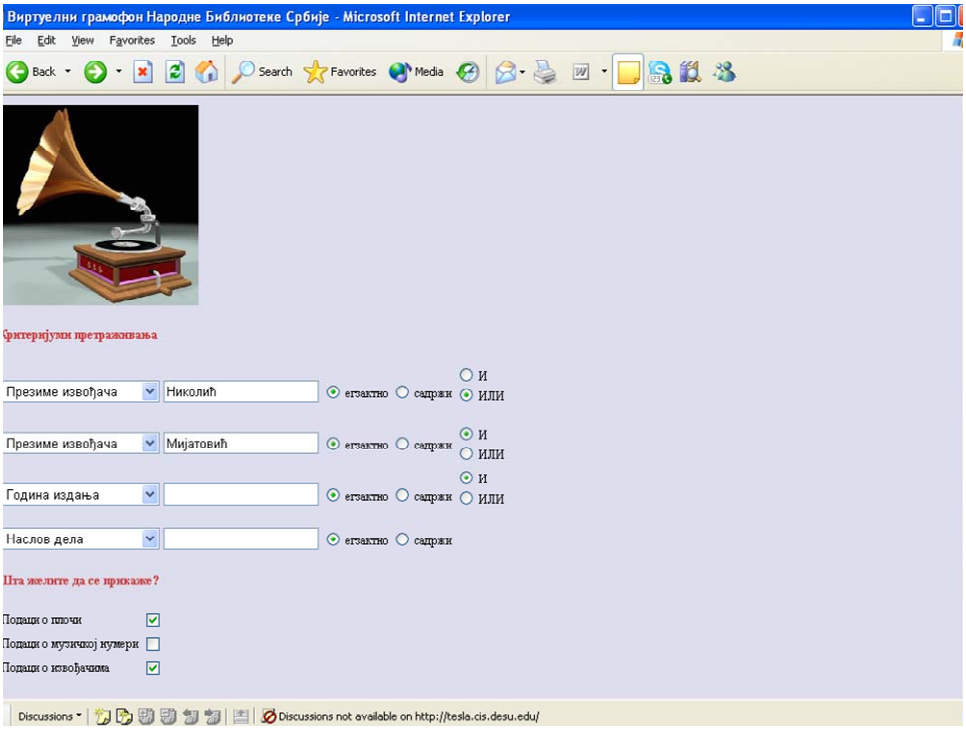
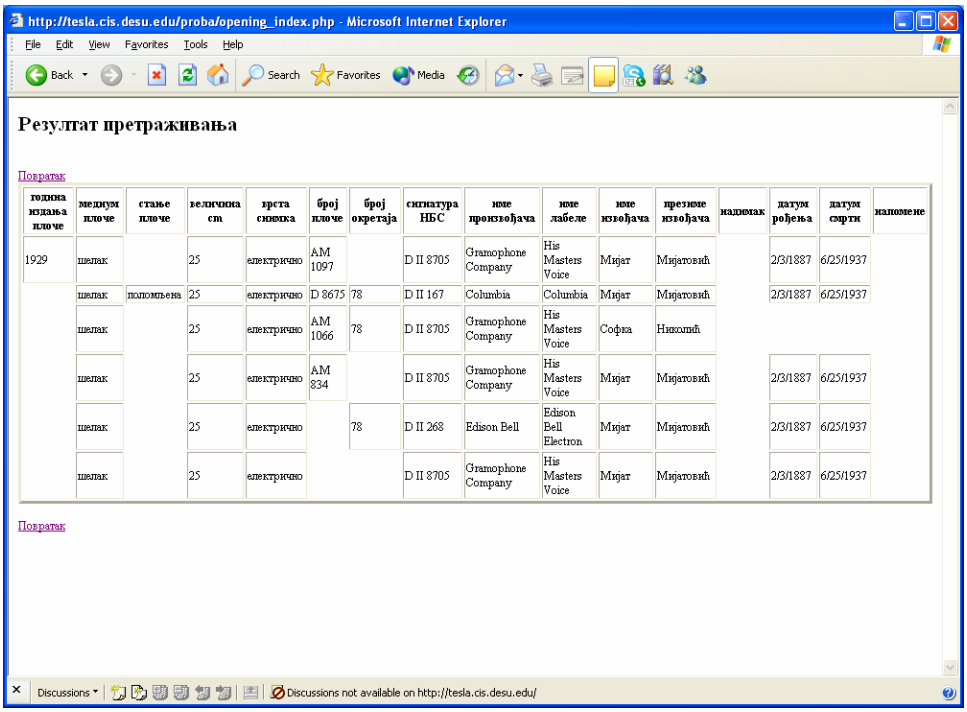


Fig. 9. An example of "or" query.



година издања плоче	медиум плоче	стање плоче	величина cm	врста снимка	број плоче	број окретаја	сигнатура НБС	име произвођача	име лабеле	име извођача	презиме извођача	надимак	датум рођења	датум смрти	напомене
1929	шелак		25	електрично	AM 1097		D II 8705	Gramophone Company	His Masters Voice	Мијат	Мијатовић		2/3/1887	6/25/1937	
	шелак	поломљена	25	електрично	D 8675	78	D II 167	Columbia	Columbia	Мијат	Мијатовић		2/3/1887	6/25/1937	
	шелак		25	електрично	AM 1066	78	D II 8705	Gramophone Company	His Masters Voice	Софка	Николић				
	шелак		25	електрично	AM 834		D II 8705	Gramophone Company	His Masters Voice	Мијат	Мијатовић		2/3/1887	6/25/1937	
	шелак		25	електрично		78	D II 268	Edison Bell	Edison Bell Electron	Мијат	Мијатовић		2/3/1887	6/25/1937	
	шелак		25	електрично			D II 8705	Gramophone Company	His Masters Voice	Мијат	Мијатовић		2/3/1887	6/25/1937	

Fig. 10. Results of "or" query from Fig. 9.

References:

[1] R. Elmasri and S. B. Navathe, *Fundamentals of Database Systems, 4th edn.*, Pearson Addison Wesley, 2003.
[2] E. Geschwinde and H. Schönig,, *PHP and PostgreSQL—Advanced Web Programming*,Sams Publishing, 2002.
[3] E. Geschwinde and H. Schönig, *PostgreSQL Developer's Handbook*, Sams Publishing, 2002.
[4] E. Kroski, *Web 2.0 for Librarians and Information Professionals*, Neal Schuman Publishers, 2008.
[5] D. Pokrajac, H. Patel, M. Rasamny, A. Sec, *Inheritance Constraints Implementation in postgreSQL*, Proc. ETRAN 48th 2004, Cacak.