

# Information Technology for Preserving the Bulgarian Folklore Heritage

Nikolay Kirov  
Computer Science Department,  
New Bulgarian University  
Montevideo Str. 21, 1218 Sofia, Bulgaria  
Email: nkirov@nbu.bg

Lozanka Peycheva  
Institute for Ethnology and Folklore Studies  
with Ethnographic Museum,  
Bulgarian Academy of Sciences, Moskovska Str. 6A,  
1000 Sofia, Bulgaria, Email: lozanka.peycheva@gmail.com

(Draft)

Abstract—Folk songs are an important and essential part of the Bulgarian cultural heritage. Following the traditions of the 20th century in publishing Bulgarian folk songs, we prepared the book “Folk Songs from Thrace” [3] with scores and lyrics recorded from original performances in the 60s and 80s of the last century. We created a digital library of over 1200 songs, which provides access to songs via full-text search engine. The data sources are stored using advanced information technology to encode texts, notes and sound. Traditional indexes and bookmarks for the book are also developed using our software.

## I. Introduction

Folk songs are an important and essential part of the Bulgarian cultural heritage. Preservation of ancient samples of Bulgarian folk music was recognized as a necessity of Bulgarian intellectual elite in the late nineteenth century: “Many educated Bulgarians in the distant and recent past realized the clear need to keep the created and preserved through centuries folk creativity in words and music” [1]. 1926 marked the beginning of systematic collection of Bulgarian folk songs with their melodies. Inspirer and organizer of this major work is Vasil Stoin. “Folk Songs from Timok to Vita” by Vasil Stoin [1] is the first capital song collection<sup>1</sup> of scores and lyrics published in Bulgaria in 1928. His project continued throughout the 20th century, he and his disciples issued 8 capital song collections (last in 1975) with a total of over 14,000 songs. Each song is presented in the books with notes (scores) and text (lyrics).

According to the method of recording the songs, collections of folk songs are of two types: recorded only as scores (using the method of “listening” or “by ear”) or recorded by audio recording devices. The songs in published capital song collections of Vasil Stoin and his collaborators (Pavel Stefanov, Ivan Kambourov, Raina Katsarova, Joseph Cheshmedjiev, Konstantin Zagorov and Hristo Iliev) [2] are collected without sound recording device. In the archive of the Institute of Art Studies there are hundreds of unpublished notated folk songs (scores and lyrics) that were written “by ear”. The songs in published capital song collections of Nikolay Kaufman,

<sup>1</sup>The term “capital song collection” („капитален песенен сборник“) is a generic name of published a large collection of folk songs (music and lyrics) from a given geographical area of Bulgaria [2].

Todor Todorov and Ilija Manolov are recorded by audio device (gramophone, tape recorder, cassette recorder, digital recorder) [2]. In the archives of the Institute of Art Studies and the Institute of Ethnology and Folklore Studies with Ethnographic Museum there is sound recorded folk music, part of which is notated, other – not.

Today, in the early twenty-first century, when we all live under the pressure of globalization and “cultural imperialism”, we feel the dilution, dispersion, scattering, attenuation of the collective memory of Bulgarian folk-musical. In the context of processes of “homogenizing” globalization and “fragmented” ethnic processes, we are increasingly aware of the need to activate the programming and development of strategies for identifying, protection, upgrading and development of folk-musical heritage of Bulgaria.

Our goal is to bring to a new level the established tradition in Bulgaria to store, organize and publish notated capital song collections of folk songs. By digitizing noted handwriting and recorded music, we aim to build a database and digital library. This allows such a structured representation of the information on certain criteria, enabling the recovery of various pieces/seams of information combined with searching by multiple criteria.

Today’s information technologies give us the possibility to digitize the existing manuscripts and musical records [6].

The source of our investigations is a collection of Todor Dzhidzhev (1927-2012), which consists of field recordings of over 1200 folk songs from the Thrace region. The archive contains:

- over 2000 pages with handwritten notes (scores) and typewriter written lyrics of songs;
- 27 tapes with recorded music tracks;
- 13 field notebooks and pads (over 500 pages) containing songs texts, notes and observations written in the time of recording the songs.

The digitization of these materials has been accompanied by many difficulties:

- technical (text recognition, coding notes, reading old magnetic types);

- identification and variations (interpretation and registration);
- 3 sources for lyrics (typewritten text, notebooks, music performances);
- obsolete and/or dialect words and pronunciation;
- grammar.

## II. Digital Library

Our database consists of files of the following types:

- $\LaTeX$  lyrics files [7]. In addition to the song text, each file contains also meta-data for the song.
- LilyPond notes files (scores). LilyPond [8] is an open source music typesetting system.
- MP3 files with authentic performance of the songs.
- JPEG files of digitized handwritten texts from notebooks and pads.

This database structure is suitable for adding new songs and new meta-data, for easy access processing programs deriving summary or classification by different criteria, for quick access of search engines by indexing, etc.

Here is an example of a  $\LaTeX$  file:

```
%td_135_1_27.txt
%Забягнала Янка робинка
%Коледарска (из пътя)
%Джиджев:коледарски

\begin{multicols}{2}
\bbf
Забягнала Янка робинка,
\uc{хей, Енке, Енке ле, Енке робинке,}
\ebf
та й бигала мало и много,
мало и много -- до девет годин,
%%чува се "до девет годин"
\five тя изляла при пашовити,
при пашовити вакли йовчари.
\kгау
Поискала йовчар да стане,
те я йотзели за йодаджия.
~~

\end{multicols}

\ed{\el{Йодаджия} -- вер. от „одая“ ’стая, гостна, килер;
спалня’; в контекста „йодаджия“ ’одалиска’}

\sing{Пял Христо Петров Петров}
\illage{гр. Тополовград}
\bornYear{1914}
\bornVillage{с. Възгаш}
\bornArea{Одринско}
\rec

%Пял Христо Петров Петров от гр. Тополовград,
%р. 1914 г. в с. Възгаш, Одринско.
%Родителите му са дошли от с. Възгаш, Одринско.
%В село Равна гора са дошли в 1916 г.
```

```
%Записал: Тодор Джиджев.
```

```
\placedate{с.\,Равна гора, общ.\,Свиленград,
11 декември 1967\,г.}
\singer{Христо Петров Петров, род. 1914 г.}
```

and LilyPond coding of the notes of the same song:

```
%{ td_135_1_27.ly }
\include "td-preamble.ly"
\score {
\relative c'' {
\tempo 4 = 168
\time 2/4
\override Glissando #'style = #'zigzag
\phrasingSlurDown
\noteFi \acciaccatura c8\glissando d4 a |
\varA
\noteFi \acciaccatura c8\startTextSpan\glissando d2
\stopTextSpan |
\varB
\afterGrace d4\startTextSpan(\( { e16[ d]\} ) c8[ b
\stopTextSpan]) |
\varC
\acciaccatura c8\startTextSpan d4 d\stopTextSpan |
\afterGrace d4(\( { e16[ d]\} ) c8[ a]) | b4 b | d8[( c) d4 |
\afterGrace d4(\( { e16[ d]\} ) c8[ b]) |
\varD
a8\startTextSpan[( b) c4\stopTextSpan |
\varE
c8\startTextSpan[( b) b( a\stopTextSpan)] |
a2 | a4 a
\bar ". " s4 \fixB e \fixC \bar ". "
\endm
\varA
e'8\startTextSpan[( f e d\stopTextSpan]) \bar "||"
\varB
d8\startTextSpan[( c] c[ b\stopTextSpan]) \bar "||"
\varC
e4\startTextSpan e8[( d\stopTextSpan]) \bar "||"
\varD
a8\startTextSpan[( b) c8 c\stopTextSpan\noBeam \bar "||"
\varE
d8\startTextSpan[( c) b( a\stopTextSpan]) \bar ". "
}
\addlyrics { За- бяг- на- ла Ян- ка ро- бин- ка,
хей, \startTextSpan Ен- ке, Ен- ке, \stopTextSpan
Ен- \startTextSpan ке ро- бин- ке \stopTextSpan }
%
\layout {
indent = #0
line-width = 190\mm
ragged-right=###f
\context {
\Lyrics
\consists "Text_spanner_engraver"
\override TextSpanner #'direction = #DOWN
\override TextSpanner #'style = #'line
\override TextSpanner #'outside-staff-priority = ###f
\override TextSpanner #'padding = #0.2
\override TextSpanner #'bound-details =
#'(left . ((Y . 0)
```

```

        (padding . 0)
        (attach-dir . ,LEFT)))
    (left-broken . ((end-on-note . #t)))
    (right . ((Y . 0)
              (padding . 0)
              (attach-dir . ,RIGHT)))) }
}
\midi { \context {
\Score tempoWholesPerMinute = #(ly:make-moment 168 4)
      }
}
}
\header{
opus = "ТД-135,1,27"
tagline = ##f
}

```

Our digital library consists of the described database and full-text search engine. The search engine provides a Google-like web interface. Table of results contains songs that match the search query. Each song is represented by a row: identifier of the song, links to all the data from the database and to the pdf files for viewing and printing of the song (scores and lyrics). For more details on the search engine, see [4]. The search engine can be used by professionals in the field of folklore research – to find common motifs, symbols or similarities between songs. These may be folk songs from different parts of Bulgaria, different versions of the same song or just general keywords.

### III. Book Preparation

Following the traditions of publication of Bulgarian folk songs in capital song collections, we prepared for printing the book “Folklore songs from Thrace” [3] based on digital library and applying new information technologies.

The book consists of 700 pages and contains 1066 songs, each song is presented by music (scores) and lyrics. The songs are arranged according to their cultural functions, because the performances of each song are associated primarily with a certain ritual, a ritual celebration or activity. The three main groups are:

- Ritual songs: calendar and family.  
[Обредни песни: календарни и семейни.]
- Labor songs: harvest, shepherd and working-bee.  
[Трудови песни: на полския труд, пастирски и седенкарски.]
- Holiday songs: dance and table.  
[Празнични песни: на хоро и на трапеза.]

The book has the following main indexes:

- Settlements:
  - Alphabetical index of settlements.
  - List of settlements in which the performers live and/or were born.
  - Settlements where there have been records of performances.

- Chronology of the records (dates and settlements).
- Performers (singers, choir and groups):
  - List of the performers with information about them.
  - List of the songs performed by each performer.
- Music bookmarks:
  - Index of the songs by time (measure).
  - Index of the songs as ornamental content.
  - Index of melodic scales as tone content.
  - Index of the songs by tonal range (ambitus).
- Vocabulary:
  - Common nouns names.
  - Anthroponyms. Toponyms. Ethnonyms. Names of holidays.
  - Numbers (numerals).
  - Glossary of dialect words.

The attached in the book CD contains the recorded “in the field” authentic performances of the songs. The disc also contains the files from the database, and the installation version of the search engine. All indexes and bookmarks from the book are included in text format, as well as the following additional applications:

- Frequency index.
- Concordance.
- Verbs. Adjectives. Adverbs.

The  $\LaTeX$  files for making the book are derived from the database with programming tools as well as indexes and bookmarks.

The main software tools are about:

- arrangement the songs and generating of  $\LaTeX$  files;
- producing the alphabetical word index in lyrics and concordance from song’s texts;
- making lists of settlements and performers;
- creating indexes of songs by time, ornament content, ambitus and sound sequences based on LilyPond files.

The software tools are written in C++ programming language and Qt as a cross-platform application and UI framework [9].

### IV. Conclusion

The collection of Todor Dzhidzhev incorporates interweaving of old items of long lasting past experience in creating capital song collections and the introduction of new elements that come in a variety of ways with tradition. The new elements are bound with the expansion of information technology in this experience.

- For the first time capital song collection is made up of pre-designed digital library.

- For the first time the issue of capital song collection is accompanied by development of software solution for searching in a digital library, according to various criteria: keywords, names of performers, names of localities and regions, etc.
- For the first time capital song collection is equipped with CD in which sound samples of collected and published songs are attached.
- For the first time verbal bookmarks are developed that navigate the availability of all (not only selected) words in the lyrics of the song collection.
- For the first time in the capital song collection are implemented: dictionary, concordance, complete list of constituent texts nouns, adjectives, numerals, verbs and adverbs.

#### Acknowledgment

This work is partially supported by Grant of the Bulgarian National Science Foundation under number DTK-02-54/2009 (see [5]).

#### References

- [1] Васил Стоин, Народни песни от Тимок до Вита, София: Издание на Министерството на Народното просвещение, 1928. [V. Stoin, Folk Songs from Timok to Vita, Published by the Ministry of National Education, Sofia, 1928.]
- [2] Списък на капиталните сборници. [List of capital song collections.] <http://nikolay.kirov.be/2010/folk/Stoin/stoin.html>
- [3] Тодор Джиджев, Народни песни от Тракия, Академично издателство „Проф. Марин Дринов“, София, 2013. [Todor Dzhidzhev, Folklore songs from Thrace, “Prof. Marin Drinov” Academic Publishing House, Sofia, 2013.]
- [4] K. Kirov, N. Kirov, Folklore Songs Digital Library and Keyword-based Search Engine, *Serdica Journal of Computing*, 6 (3), (2012), 163-174.
- [5] Information technologies for presentation of Bulgarian folk songs with music, notes and text in a digital library. <http://nikolay.kirov.be/2010/folk>
- [6] L. Peycheva, N. Kirov, Bulgarian Folk Songs in a Digital Library, In: Proc. Int. Conf. Digital Preservation and Presentation of Cultural and Scientific Heritage, 11-14 September 2011, Veliko Tarnovo, Eds. R. Pavlov, P. Stanchev, (2011), 60-68.
- [7] L<sup>A</sup>T<sub>E</sub>X typesetting system. <http://www.latex-project.org/>
- [8] Lilypond music engraving program. <http://lilypond.org/>
- [9] Qt Project, <http://qt-project.org/>