# Research on Linguistic Approaches, Used for Semantic Explanation of Bell's Knowledge

Galina Bogdanova[1], Todor Todorov[1], Nikolay Noev[1], Stefka Kancheva[2]

[1]Institute of Mathematics and Informatics, BAS
{galina,todor}@math.bas.bg, nickey.noev@gmail.com
[2]University of Veliko Tarnovo "St. St. Cyril and Methodius"
s_kancheva@uni-vt.bg

**Abstract.** This paper presents a research of linguistic structure of Bulgarian bells knowledge. The idea of building semantic structure of Bulgarian bells appeared during the "Multimedia fund – BellKnow" project. In this project was collected a lots of data about bells, their structure, history, technical data, etc. This is the first attempt for computation linguistic explain of bell knowledge and deliver a semantic representation of that knowledge. Based on this research some linguistic components, aiming to realize different types of analysis of text objects are implemented in term dictionaries. Thus, we lay the foundation of the linguistic analysis services in these digital dictionaries aiding the research of kinds, number and frequency of the lexical units that constitute various bell objects.

**Keywords:** Multimedia Digital Libraries, Frequency Dictionaries, Bells, Cultural Heritage, Multimedia Digital Archives, Computer Science.

## 1 Introduction

The idea of the building of the semantic explanation of the collected knowledge about bells was born during the realization of BellKnow project, which aim was to collect different type of data about bells. Based on this research some linguistic components are implemented in term dictionaries. Thus, we lay the foundation of the linguistic analysis services in these digital dictionaries aiding the research of kinds, number and frequency of the lexical units that constitute various bell objects. During the process of design of Bell ontology we created different types of dictionaries which we will present in this paper. This is the first attempt for computation linguistic explain of bell knowledge and deliver a semantic representation of that knowledge.

## 2 Research Background

During the project "Multimedia fund – BellKnow"[1] was collected digital multimedia data about 120 bells, situated in 20 sites (churches, bell towers, museums).

The establishing of Multimedia fund "BellKnow" was meant to build a multimedia depository of the records and research results which we did during the years.

At first BellKnow project gathered information about every particular bell. It was made different examinations from different area's experts from various institutions. It was made following examinations: taking the dimensions of of the shape and geometry of the bells; photographing of the outward appearance and the ornaments of the bells with a professional equipment; recording the sound of the bells with a professional equipment and special microphone; analyzing bell's sound, including sound frequency, fade curve line, stroking tone and other parameters of the sound [9]; measuring the exact location of every particular bell; recording the stroke of every particular bell with a camera; studying the origin of the bell; estimate the condition of the bell; studying of art and cultural value of the bell; etc.

The next step was making of passport of every particular bell, filled with the data of the previous investigation. In result the multimedia fund "BellKnow" was filled with content more than 3 000 bell's digital records, described in [1, 4, 7, 10].

During the work on "BellKnow" we exchange and share information on specific acoustic investigations on bell sound with a Russian scientists [15].

We share our experience with colleagues from various institutions and made together researches in the area of multimedia funds in different fields of knowledge [5, 12, 13]. We had a project on a field of digitalization of folklore heritage with Institute of Ethnology and Folklore Studies with Ethnographic Museum and University of Veliko Tarnovo [2, 3]. Also we made researches in a building of information system for linguistic data processing [6, 11, 14]. We exchange experience with Swiss scientists in the field of semantic interpretation of 3D pint clouds [8].

## 3 Semantic Concept of Bulgarian Bells

### 3.1 Semantic Web

The main subject of the idea of Semantic Web, offered by Tim Berners-Lee consist in automation of "intelligent" processing of knowledge of a different Internet resources in semantic meaning or is an explanation and transformation of knowledge in machine-interpretable definitions, through which the computer semantic agents could draw conclusions.

For design of a semantic in particular area mostly is used W3C Semantic Web initiative standards, which are SGML, XML, RDF, OWL [16] etc. They are used for the

[1] The "Multimedia fund – BellKnow" is research project of the Institute of Mathematics and Informatics. Its main goals are to build a multimedia digital fund with a set of various objects/collections, selected from the Multimedia digital archive of Bulgarian Bells.

description of the information and add an additional meaning, which could be used by semantic search agents.

We make an experimental semantic annotation, based on the current W3C Semantic Web initiative standards (RDF, RDFS and OWL, etc.) of the resources in digital archive of unique bells. We use the RDF data model, because it provides a model for describing resources of bells. Digital resources have properties (attributes or characteristics of bells). RDF defines a digital resource as any object that is uniquely identifiable by a Universal Resource Identifier (URI).

### 3.2 Research on Linguistic Explanation of Bell's Knowledge

The linguistic explanation is very essential task in our days. Basically the semantic systems are built of different types of thematic dictionaries, which explain nature (give a meaning) of the information, preserve in data banks.

Numerous ontology dictionaries with different functions are a basis of our experimental bell ontology system. These dictionaries are made using XML/RDF/OWL technology to provide lexical explain and description of bell knowledge. Combination of descriptions, text explanation and lexical rules provides basics of experimental bell ontology system.

### 3.3 Bell Ontology

With collected knowledge during the project Multimedia fund BellKnow, we built an experimental Bell ontology, in which we gave meaning of the different data and established connections between the different concepts. We design the single elements and data in separate objects and logic connected them with other objects. With that experimental practical ontology or more exactly semantic taxonomy of bell's knowledge, we tried to connect semantic objects-bells with different knowledge such as places, historical events, technical data, production technology, sound measurements and etc. [6].

## 4 Linguistic Approaches of Semantic Research

The main tools to represent semantics are dictionaries, which concepts are arranged in a subtype-supertype hierarchy, thus forming taxonomy.

### 4.1 Frequency Dictionary for Bell Science

The main purpose of this project is to build a text-research system to create frequency dictionary from texts of bell knowledge descriptions. To construct this system are built a hierarchical datasets and WEB interface [3, 11]. The texts are divided into thematic headings (bell science domains).

The dictionary aims to cover big corpuses of texts that contain terms and conceptions in the field of Bulgarian bell's knowledge. Frequency dictionary is designed for

large arrays of such texts. Separate phrases and words are representatives of different domains. The dictionary aims to give information on how often a particular word or phrase is used in a particular corpus of texts.

Full text search could be made with the system. After the indexing of the data, searching of the appropriate index is saved in catalog-table. A problem occurs when there is a need of countdown of particular phrase when a particular criterion is used. In result is shown the information in how many files in one particular area's domain, how many words and phrases are founded. A hierarchical structure of data (tree) is used for organization of the information. The hierarchical structure of data has tables included for administration of categories and growing of the tree structure is allowed in volume and depth.

The system offers an easy and fast search system, due to the hierarchy of the data. It enables introduction of many different domains and nevertheless they don't influence the speed of searching. The individual tables contain only the names of domains, as well as their keys for organization the hierarchy. The help table contains all texts of all domains, organized with the help of indexes, which enables a fast access to the relevant texts and domains. There is an option for construction of a dynamic growing of the tree of tables in depth.

### 4.2 Term Dictionaries for Bell Science

For the bell science domain more suitable dictionaries are term dictionaries. The term dictionaries present the explanation of bell knowledge in a definite corpus of texts. It is considered that the facts in term dictionary are enough. The term dictionaries give versatile information: text context of different terms, presence/absence of define terms, relations with other terms, lexical rules, groups of different bell domains.

Further is shown a part of XML dictionary data:

```
<bdict>
  <item>
    <idbt>1</idbt>
    <term lang="bg">камбана</term>
    <description lang="bg">музикално
устройство</description>
    <type lang="bg">съществително</type>
    <classOf>
      <clOf lang="bg">предмет</clOf>
      <clOf lang="bg">музикално устройство</clOf>
    </classOf>
    <ref>
      <href lang="bg">камбана</href>
      <href lang="bg">звънец</href>
      <href lang="bg">хлопатар</href>
    </ref>
    <transl>
```

```
      <bword lang="en">bell</bword>
    </transl>
    <bnote lang="bg">камбана</bnote>
  </item>
…
</bdict>
```

**Bell Termin Dictionary**

main | search word | all of termins | advanced | show xml

advanced | add item | edit termins | detete termins | edit translate | simple xml object | show xml

**List of Bell Termins**

| ID | termin | description | type | class | href | translate | note | edit |
|---|---|---|---|---|---|---|---|---|
| 1 | камбана | музикално устройство | съществително | предмет, музикално устройство | камбана, звънец, хлопатар | bell | камбана | edit |
| 2 | камбанария | сграда в която се помещава камбана | съществително | предмет, музикално устройство | камбана, църковна камбанария | belfry, belfry 2 | камбанария | edit |
| 3 | звънене | звънене | глагол | предмет, музикално устройство | дрънчене | sound | звънене | edit |

**Fig. 1.** Bell Term Dictionary v1.1

The final line we looking for is to build a semantic ontology system, that contains a huge amount of data and linguistic explain of bell knowledge.

## References


1. Bogdanova G., Dimkov G., Todorov T., Noev N.: Model of Digital Repository of Information and Knowledge - Fund “BellKnow”, Annual Seminar on algebraic and combinatorial coding theory, Ksilifor, Bulgaria (2010)
2. Bogdanova, G., Pavlov, R., Todorov, G., Mateeva, V.: Technologies for Creation of Digital Presentation and Significant Repositories of Folklore Heritage. Advances in Bulgarian Science Knowledge, National Center for Information and Documentation, 3, 7–15 (2006)
3. Bogdanova G., Rangochev K., Paneva-Marinova D., Noev N.: Towards Linguistics Analysis of the Bulgarian Folklore Domain, International Journal "Information Technologies and Knowledge", Vol.5, №2, pp. 119–128, ISSN 1313-0455, presented at International conference - i.Tech’11, Varna, Bulgaria (2011)
4. Bogdanova G., Todorov T., Noev N.: Digital Repository of Information and Knowledge - Fund "BellKnow", In the Proceedings of the First International Conference “Digital Preservation and Presentation of Cultural and Scientific Heritage” - DiPP`11, pp. 91-98, ISSN: 1314–4006, Veliko Tarnovo, Bulgaria (2011)
5. Bogdanova G., Todorov T., Noev N.: Digitalization and security of "Bulgarian Folklore Heritage" archive, CompSysTech, ACM International Conference Proceeding Series (ICPS) vol. 471, pp. 335–340, (2010)

6. Bogdanova G., Todorov T., Noev N.: Singing individual fragments of an RDF graph of unique Bulgarian bells, In: Twelfth international workshop, ACCT`2010, ISBN 978-5-86134-174-5, pp. 47–52, Academgorodok, Novosibirsk, Russia (2010)
7. Bogdanova G., Trifonov T., Todorov T., Georgieva T.: Methods for Investigation and Security of the Audio and Video Archive for Unique Bulgarian Bells. National Workshop on Coding Theory and Applications, Blagoevgrad (2006)
8. Cotofrei P., Künzi Ch., Stoffel K.: Semantic Interpretation of 3D Point Clouds of Historical Objects. In: 1st International Conference “Digital Preservation and Presentation of Cultural and Scientific Heritage” - DiPP`11, pp. 127–139, ISSN: 1314-4006, Veliko Tarnovo, Bulgaria (2011)
9. Dimkov G., Alexiev Al., Simeonov I., Trifonov T., Simeonov K.: (in Bulgarian) Acoustic research of historic value Bulgarian bells, National Conference Akustika'2008, 116 ISSN 1312–4897, Varna, Bulgaria (2008)
10. Noev N.: Organization and Security of the Audio and Video Archive for Unique Bulgarian Bells. Mathematica Balkanica, NewSeries Vol. 24, ISSN: 0205-3217, Fasc.3–4, pp. 285–291 (2010)
11. Noev N., Hristov G., Stanchev I.: (in Bulgarian) "Experimental specialized frequency dictionary," National Conference "Innovations in software technology, algorithms and learning in higher education and related", pp. 102–109, Apriltsi, Bulgaria (2009)
12. Paneva-Marinova, D., Pavlov R., Rangochev K.: Digital Library for Bulgarian Traditional Culture and Folklore, In: 3rd International Conference dedicated on Digital Heritage (EuroMed 2010), pp. 167–172, Published by ARCHAEOLINGUA, Lymassol, Cyprus (2010)
13. Pavlov R., Paneva, D.: Toward Ubiquitous Learning Application of Digital Libraries with Multimedia Content. Cybernetics and Information Technologies, Vol. 6, № 3, pp. 51–62 (2007)
14. Pavlov R., Pavlova-Draganova L., Draganov L., Paneva, D.: e-Presentation of East-Christian Icon Art. In: Fourth HUBUSKA Open Workshop “Semantic Web and Knowledge Technologies Applications”, Varna, Bulgaria, pp. 42–48 (2006)
15. Sapozhnikova K., Taymanov R.: Acoustic Riddles of Cultural Heritage. In: 1st International Conference “Digital Preservation and Presentation of Cultural and Scientific Heritage” - DiPP`11, pp. 79–90, ISSN: 1314-4006, Veliko Tarnovo, Bulgaria (2011)
16. Web Ontology Language, http://www.w3.org/TR/owl-features/