

Ученые записки Таврического национального университета им. В.И. Вернадского  
Серия "Филология". Том 20 (59). 2007 г. №4. С.68-71.

УДК 81'374=161.1 004.415

## КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ СОЗДАНИЯ ЭЛЕКТРОННОГО СЛОВАРЯ

*Чепик Е.Ю*

*Таврический национальный университет им. В.И. Вернадского,  
г. Симферополь, Украина*

*В статье рассматривается модель электронного словаря политической лексики, приводится общая концепция словаря и его структуры, описывается программная оболочка и перспективы дальнейшей разработки*

*Ключевые слова:* компьютерная лексикография, электронный словарь

*Актуальность.* Развитие информационных технологий является одним из знаковых явлений начала XXI в. Использование достижений таких технологий в различных областях науки представляется сегодня важной составляющей различных исследований. Так, в современной лингвистике информационные технологии используются для автоматической обработки текстовой информации, представленной на естественных языках. «Язык, являясь универсальным средством представления информации, в то же время не является по своей сути формальным в той мере, в какой он мог бы пониматься и интерпретироваться компьютером, являющимся основным звеном в современных информационных технологиях» [7, с. 3]. Одним из способов формализации языка является составление лексикографического массива, а именно создание отраслевых словарей, охватывающих только одну тематическую область. Этим и обусловлена актуальность предлагаемого исследования. В современной компьютерной лингвистике важной является проблема автоматизации составления, ведения и лингвистического обеспечения электронных словарей [7; 1; 6; 3; 5]. Организация баз данных, на основе которых проводится формирование словарных статей, способы презентации материала в гипертекстовом пространстве словаря, обоснование лингвистического обеспечения процессов поиска информации в базах данных определяют научную новизну работы.

В наших работах [8; 2] был представлен подход к реализации электронного словаря политической лексики, а также достаточно полно описана модель словаря, которая может быть использована для реализации программных продуктов, ориентированных на другие, отличные от политической, области знаний. В качестве объекта исследования была выбрана политическая лексика именно потому, что она наиболее подвержена семантическим изменениям, и это особенно заметно в периоды политических и социальных изменений. Современный политический язык активно меняется под влиянием экстралингвистических факторов и, таким образом, требует репрезентации в часто обновляемом продукте, соответствующем современным реалиям. Цель статьи – рассмотреть модель электронного словаря политической лексики.

Программная оболочка словаря написана на языке программирования Delphi. Она состоит из двух частей: первая формирует запрос и выводит результаты, вторая обеспечивает обработку запроса, взаимодействие с базой данных в Access.

Пользователь видит при запуске словаря только первую часть программной оболочки, то есть ее интерфейс, позволяющий визуализировать все необходимые пометы и другие элементы оформления (курсив, подчеркивание, отступы, цвет текста и т.д.), присутствующие в обычном (бумажном) словаре. В карточки словаря также могут быть включены ссылки на файлы-иллюстрации, перекрестные ссылки между несколькими карточками, ссылки на литературу.

Таким образом, автоматизация лексикографической работы, под которой мы понимаем технологическую линию, где на разных этапах применяются разные пакеты программ – как специализированные, так и общие, которые комплексно внедряют в лексикографию новейшие информационные технологии [4], в данном лексикографическом продукте позволяет осуществлять задачи, связанные с автоматическим поиском слов, добавлением слов в базу данных, озвучиванием заголовочного слова, экспортированием в HTML, обработкой информации в пользовательском интерфейсе, а также удобной навигацией внутри словаря. Кроме этого, пакет программ позволяет расширять возможности лексикографического продукта.

Доступность словника или всех словарных статей для пользователя обеспечивает поле поиска и последовательный просмотр словника или словарных статей с помощью «ролистывания» («рокрукти») слов в поле быстрого поиска.

Лингвистическая база данных электронного словаря политической лексики (ЛБД), т.е. «пециальная структура данных, которая служит для хранения и эффективного использования исходной, промежуточной и конечной информации о ЕЯ объектах» [7, с. 75 -76], в рассматриваемом словаре политической лексики включает в себе информацию, разделенную на 8 структурных компонентов, каждый из которых представлен в виде столбца в базе данных Access:

- 1) *ID* – уникальный идентификатор статьи;
- 2) *word* – заголовочное слово;
- 3) *word author* – авторское толкование;
- 4) *field\_illust* - иллюстрация употребления слова;
- 5) *field\_expl\_dict* – данные из толковых словарей;
- 6) *field\_encyclp\_info* – данные из энциклопедических словарей;
- 7) *sound* - звуковой файл для соответствующей лексемы;
- 8) *field\_history\_word* – данные об истории слова.

id	word	word_author	field_illust	field_expl_dict	field_encyclp_info	sound	field_history_word
3	Авторитария	Авторитария, и Представители Бакеркина В.Е. Большой энциклопедический словарь					
5	Амбиции	Амбиции, ий, Если раньше в Бакеркина В.Е. (http://encyclp.info) и другие данные					
6	Битники	Битники, ов, Кан и поллюда Бакеркина В.Е.					
8	Брифинг	Брифинг, а в 1 июля право Бакеркина В.Е. Большой энциклопедический словарь					
10	Бюрократия	Бюрократия Иерархическая Бакеркина В.Е. Большой энциклопедический словарь					
13	Вето	Вето, несл., На последнем Бакеркина В.Е. (http://encyclp.info) и другие данные					1370 1975 - 3 доку

Рис.1. Образец лингвистической базы данных словаря политической лексики

В конечном лексикографическом продукте структурные компоненты 2, 3, 4, 5, 6, 8 (см. Рис.2) репрезентируются в отдельных полях, визуальнo представленных в виде окон, а столбец 7 представлен в виде кнопки, которая вызывает воспроизведение звукового файла из базы данных.

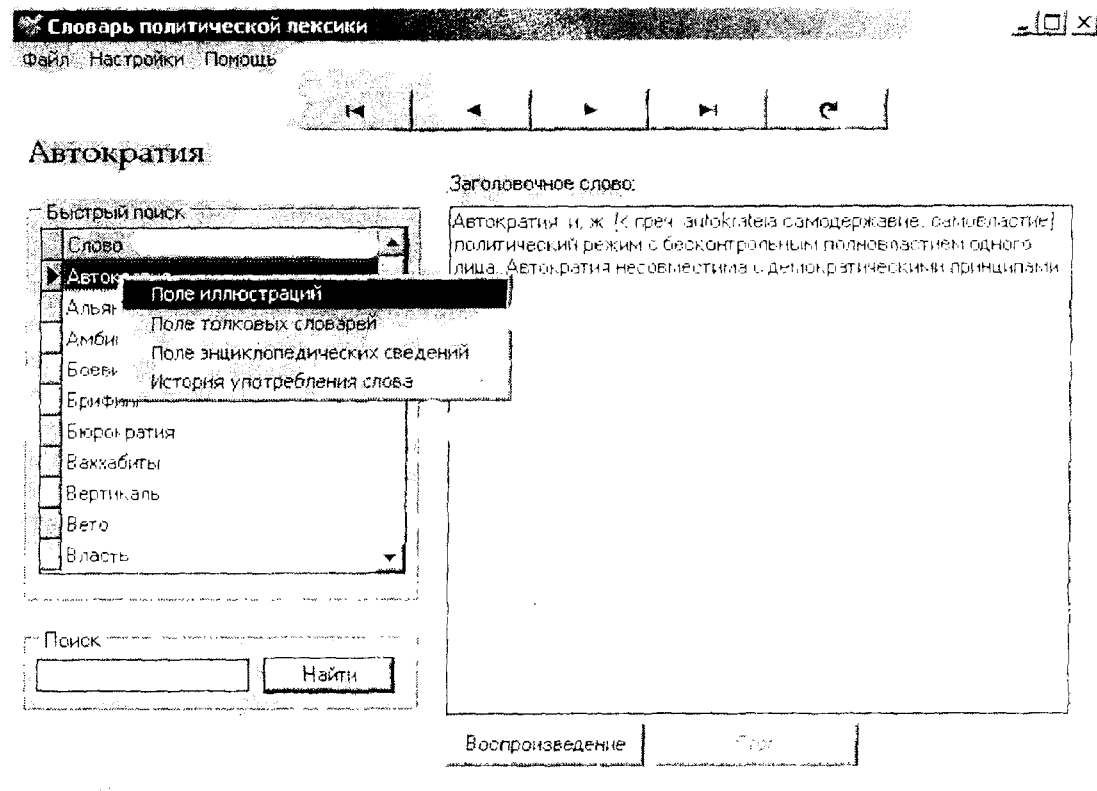


Рис.2. Образец конечного лексикографического продукта

В перспективе данного исследования планируется введение «поля нового слова», позволяющего получить информацию по 4-м полям о дефиниции, которой нет в словаре. Для этого ставится задача написания программного обеспечения, способного осуществлять сканирование по базе из 10 словарей, а также из Русского национального корпуса и выводить сведения по каждому полю отдельно. Эта функция даст возможность оперативно вводить дефиниции, которые действительно отражают современную языковую ситуацию.

*Выводы.* В заключение следует отметить, что методы автоматизации ведения и составления словарей не унифицированы. Это позволяет находить новые подходы к отражению предметной области, рассматриваемой в словаре.

### Список литературы

1. Белоногов Г.Г., Калинин Ю.П., Хорошилов А.А. Компьютерная лингвистика и перспективные информационные технологии – М.: Русский мир, 2004 – 248 с.
2. Бессонова Л.Е., Чепик Е.Ю. Опыт создания электронного словаря политической лексики: лингвистическая модель и программная оболочка/ MegaLing 2006. Горизонты прикладной лингвистики та лінгвістичних технологій//Доповіді міжнародної конференції.20-27 вересня 2006. Україна, Крим, Партеніт/Ред.В.А. Широков, С.С. Дикарева. – Сімферополь: Вид-во «ДиАйПи», 2006. – С. 139-140.

3. Волкова И.А. Введение в компьютерную лингвистику. Практические аспекты создания лингвистических процессоров: Учебное пособие для студентов факультетат ВМиК МГУ. – М.: Издательский отдел факультетат ВМиК МГУ им. М.В. Ломоносова, 2006. – 44 с.
4. Кульчицький І.М. Комп'ютерно-технологічні аспекти створення сучасних лексикографічних систем. – К.: Національна бібліотека України ім. В.І.Вернадського НАН України, 2002. – 57 с.
5. Проблемы компьютерной лингвистики: Сб-к научн. трудов/ Под ред. А.А. Крегова. – Вып. 2. – Воронеж, 2005. – 210 с.
6. Филиппович Ю.Н. Метафоры информационных технологий: рабочие материалы исследования. С предисловием Ю.Н. Караулова. – М.: МГУП, 2001. – 288 с.
7. Филиппович Ю.Н., Прохоров А.В. Семантика информационных технологий: Опыт словарно-тезаурусного описания. С предисловием А.И.Новикова. – М.: МГУП, 2002. – 368 с.
8. Чепик Е.Ю. Политическое слово в структуре электронного словаря // Культура народов Причерноморья. 2005. – №69. – С. 205-210.

**Чепик О.Ю. КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ СТВОРЕННЯ  
ЕЛЕКТРОННИХ СЛОВНИКІВ**

*У статті розглядається модель електронного словника політичної лексики, наводиться загальна концепція словника і його структури, описується програмна оболонка й перспективи подальшої розробки.*

*Ключові слова:* комп'ютерна лексикографія, електронний словник

**Chepik E.U. ELECTRONIC DICTIONAR: PROGRAM ENVIRONMENT  
AND PROSPECTS OF THE FURTHER DEVELOPMENT IS DESCRIBED**

*In article the model of the electronic dictionary of political lexicon is considered, the general concept of the dictionary and its structure is resulted, the program environment and prospects of the further development is described*

*Key words:* a computer lexicography, the electronic dictionary

*Поступили в редакцію 19.02.2007 г.*