

УДК 371.64/.69:004

ЭТАПЫ ПРОЕКТИРОВАНИЯ ЭЛЕКТРОННЫХ СЛОВАРЕЙ

Балалаева Елена Юрьевна

Национальный университет биоресурсов и природопользования Украины,
г. Киев, ул. Героев Оборона, 15; email: olena.balalaeva@gmail.com

***Аннотация:** В статье проанализированы современные подходы к процессуальному аспекту проектирования электронных словарей в теории и практике компьютерной лексикографии. На основе анализа научных источников, сопоставления и обобщения опыта ученых и собственных исследований предложено авторское видение вопроса проектирования электронного словаря как процесса. Приведены характеристики аналитического, концептуального, технологического, экспериментального и рефлексивного этапов проектирования электронного словаря.*

***Ключевые слова:** электронный словарь, проектирование, процесс, этап, процедуры*

Актуальность. По мнению многих исследователей, новая информационная парадигма требует переоценки стратегий составления лексикографических источников, содержания информации и способов взаимодействия с новым типом пользователя – «digital native» [1]. В фокусе внимания оказываются три аспекта: информация, доступ к ней и пользователь. Содержание информации разрабатывает лексикография, доступ к ней – компьютерные науки, построение моделей взаимодействия с пользователем – социология и психология. Все это актуализирует исследования проблемы проектирования электронных словарей в теоретическом и практическом аспектах, ведь этот вид лексикографических ресурсов является весьма популярным среди пользователей [2].

Данная статья сосредоточена на процессуальном аспекте проектирования электронного словаря, предполагающем разработку и реализацию этапов проектной деятельности.

Описание процесса создания электронных словарей можно найти в трудах многих ученых, в частности Н. Агаповой, Л. Беляевой, А. Ефремовой, Е. Карпиловской, Н. Картофелевой, М. Ковязиной, В. Табанаковой, С. Шереметьева и др. Однако следует согласиться с мнением И. Кудашева, что ни одно из этих описаний отдельно, ни их совокупность не могут претендовать на роль полной и универсальной схемы этапов работы над словарем [3, с. 50]. Поскольку процесс создания электронных словарей динамично развивается, он требует регулярного пересмотра, систематизации и рефлексии в научной литературе.

Цель статьи – на основе анализа современных подходов к процессу создания электронных словарей, дать характеристику этапов их проектирования. В исследовании применены методы анализа научных источников, сравнения, обобщения и систематизации.

Результаты исследования. Проектирование электронного словаря в этом исследовании трактуется как процесс создания научно обоснованного проекта электронного словаря с заданными характеристиками, свойствами, параметрами и функциями, детерминирующими его макро- и микроструктуру, который не может быть трансформирован в бумажный аналог без редукции его функций.

В конструкторских дисциплинах проектирование как процесс, разворачивающийся во времени, разделяют на стадии, этапы, процедуры и операции: стадии состоят из объединенных характером работ этапов, которые, в свою очередь, состоят из определенных процедур – формализованной совокупности действий или операций [4].

В лексикографических исследованиях такого детального подхода придерживаются не всегда и чаще всего оперируют понятием «этап» в широком смысле. Более того, различные

исследователи имеют разные представления не только о характеристиках конкретных этапов, но и о границах самого процесса проектирования словаря.

Так, И. Кудашев условными границами этапа проектирования считает промежуток от формулировки общего замысла словаря до оформления результатов проектирования в виде концепции [3, с. 51].

Исследовательницы В. Табанакова, М. Ковязина считают, что моделирование электронного словаря начинается с определения и обоснования его типа, типологических параметров и разработки комплекса соответствующих лексикографических средств (метаязыка словаря) [5].

Среди факторов, влияющих на параметры словаря и процесс его создания, различают собственно лексикографические и внешние. Первые связаны с удовлетворением потребностей пользователей словаря, вторые – корректируют проект словаря соответствии с реальными условиями [3, с. 66-70].

К собственно лексикографическим относят следующие факторы: пользователи, материал, с которым оперируют пользователи, характер этих операций, а также производные от суммы указанных факторов информационные потребности пользователей и потребности, связанные с особенностями поиска и восприятия информации, к внешним факторам относят ресурсы и ограничения, обусловленные носителем информации.

Факторами, влияющими на выбор предметной области словаря, являются: социальный заказ, уровень предметной компетенции составителя, наличие доступа к материалам и специфика выбранной отрасли.

Как отмечает В. Широков, важным понятием для определения структуры словарных систем является введенное Ю. Карауловым понятие лексикографического параметра – определенного «кванта» лингвистической информации, которая может иметь самостоятельный интерес для пользователя, но, которая, как правило, выступает в комбинации с другими параметрами и находит свое специфическое выражение в словарях. Лексикографическим параметрам присущи определенные свойства, которые имеют универсальный характер и не зависят от типа словаря, а их значения характеризуются определенной спецификой конкретного языка. Количество лексикографических параметров в словарях варьируется от одного до нескольких десятков, причем эти колебания определяются не только целевой направленностью словаря, но и его ориентацией на пользователя [6, с. 13-15].

На этапе разработки общей концепции словаря целесообразно различать пользователей в соответствии с их информационными потребностями, являющимися производными от уровня их компетенции, а именно: предметной, лингвистической и лексикографической [3, с. 120].

По мнению многих исследователей, словари в электронной форме, адресованные человеку, отличаются от бумажных словарей наличием в них технического инструмента, который позволяет предоставлять пользователю словарный контент, релевантный его запросу: морфологический и синтаксический анализ, полнотекстовый поиск, гипертекст, распознавание и синтез звука, а также позволяет предоставлять этот содержание разными способами: аудио, графические средства и т.д. [7].

Большинство исследователей принципиальными положительными отличиями электронных словарей считают их объем, нелинейную структуру и расширенные возможности поиска. Электронные имеют более широкие возможности по поиску информации. Если в бумажных словарях вход, как правило, возможен только через форму единиц словаря или через несколько параметров (при наличии дополнительных указателей), то в электронных словарях появляется возможность отбирать статьи по любым критериям и их комбинациям.

По мнению В. Пастор и А. Альцины, все параметры в электронных словарях определяют три элемента: запрос, ресурс и результат [8, с. 309]. Типы запросов подпадают под диапазон типов поиска, среди которых исследовательницы различают: поиск,

аналогичный алфавитному в бумажных словарях; вспомогательный поиск, при котором словарь предлагает слова в случае неправильного ввода пользователем; расширенный поиск, который позволяет пользователю искать в словарном контенте определения, примеры и т. п.; поиск с символами подстановки, шаблонами и фильтрами, позволяющий сочетать поисковые слова с операторами и осуществлять поиск орфографических вариантов с символами подстановки; поиск родственных слов; поиск по анаграммам, аббревиатурам и сокращениям; усовершенствованный поиск (по произношению или этимологии). Ресурсы, содержащие поисковую информацию, делятся на четыре типа: поле ввода, поле контента, индекс тематического поля и поле доступа к внешним ссылкам. Результатом поиска является информация, которую получает пользователь в ответ на запрос. Обычно это запись информации о слове (значение, грамматическая информация, произношение, этимология, использование в контексте эквивалентность и т.д.) или слово или список слов, соответствующих записи в словаре. Также результатом может быть изображение или перечень изображений, аудиофайлы. Получение этих результатов зависит от типа словаря и его. Именно три элемента: запрос, ресурс и результат детерминируют параметры электронных словарей [8, с. 352].

Поэтому в проектировании словарей, ориентированных на человека как конечного пользователя, следует учитывать и некоторые неинформационные потребности (эргономичную, экономическую и эстетическую) [3, с. 121].

Н. Агапова, Н. Картофелева [9] выделяют шесть этапов создания электронного словаря:

1. Сбор материала, написание словарных статей.
2. Разработка общей концепции словаря, выбор формы существования словаря, составление программистом технического задания.
3. Выявление функциональных и нефункциональных требований, детализация технического задания на основе выявленных требований, написание словарных статей для электронного словаря.
4. Проектирование архитектуры компьютерной программы с несколькими вариантами поиска и группировки материала, выбор средств разработки.
5. Реализация электронного словаря на основе спроектированной архитектуры приложения; разработка программной платформы словаря, создание интуитивно понятного интерфейса и удобной навигации.
6. Апробация созданного электронного словаря.

С. Шереметьева и др. [10] также описывают шесть этапов создания электронного словаря, но с другими задачами, операциями и процедурами. На первом этапе определяют предметную область, цель и пользователей словаря, на втором – модель знаний: типы, объем и формализм подачи лингвистической информации в словарной статье. Третий этап заключается в определении источника и состава вокабуляра в соответствии с моделью знаний. На четвертом этапе определяют эквиваленты для каждой единицы вокабуляра с лингвистической информацией о переводных эквивалентах в соответствии с моделью знаний. Пятый этап предполагает спецификацию и программную реализацию словаря, шестой – внесение информации в электронную словарную оболочку.

А. Ефремова [11] предлагает пять этапов создания электронного словаря:

1. Выбор исходных лексикографических ресурсов.
2. Разработка макро- и микроструктуры электронного словаря.
3. Проектирование базы данных электронного словаря.
4. Распознавание макро- и микроструктуры выходных словарей и заполнения базы данных электронного словаря.
5. Разработка программного обеспечения для работы с базой данных словаря и его тестирование.

С. Загайнов [12] рассматривает четыре этапа проектирования электронного словаря:

1. Аналитический этап, во время которого обосновывают актуальность проекта, определяют перспективные для проекта технологии и программы, формулируют концепцию электронного словаря.

2. Подготовительный этап предполагает определение основных принципов работы электронного словаря и разработки его общей структуры с опорой на результаты первого этапа исследования, анализ возможных лексикографических источников и источников экстралингвистической информации.

3. Практико-технический этап направлен на создание экспериментальной словарной базы, определения связей между словарными статьями для разработки системы ссылок между ними.

4. Экспериментальный этап заключается в проверке выполнения прототипом электронного словаря гипотезы исследования, проведение эксперимента, направленного на определение эффективности разработанной концепции.

С одной стороны, электронный словарь является лексикографическим объектом, поэтому при его проектировании можно ориентироваться на общие принципы описания лексикографических единиц, среди которых выделяют следующие:

- ориентация на адресата;
- множественность лингвистического описания (существует множество моделей описания лингвистического материала, не исключая друг друга);
- стандартность (явления одного порядка в одном словаре требует унифицированного представления);
- экономичность (при прочих равных условиях предпочтение отдается максимально лаконичному описанию);
- полнота (отражает необходимость исчерпывающего лексикографического представления всех существенных признаков значения и правил употребления языкового знака);
- эффективность (обеспечивающая некую алгоритмизацию словарной статьи с тем, чтобы прочтение дефиниции давало возможность точного понимания описываемого языкового факта в широком контексте с учетом всех парадигматических, синтагматических и других свойств знака);
- семантическая ступенчатость (определяет выбор метаязыка словарного описания таким образом, чтобы его элементы были только на один признак проще/сложнее определяемого понятия) [14].

С другой стороны, проектирование электронных словарей осуществляют с учетом специфических принципов, к которым Л. Беляева относит следующие:

- модульность (словарные статьи проектирует как относительно независимые объекты, организованные в системы, которые, в свою очередь, рассматривают как подсистемы системы обработки данных);
- динамичность (возможность оперативно корректировать номенклатуру словаря, состав и объем информации словарной статьи, а также вносить в систему новые программы и файлы);
- гибкость (способность баз данных и словарей соответствовать новым требованиям без кардинальной реструктуризации);
- сбалансированность (определяет системность комплектования баз данных и словарей, исключает произвольность и стихийность в отборе информации);
- дружелюбность (обеспечивает эффективный характер взаимодействия между системой и пользователем и удобство пользования) [7, с. 13].

В создании электронного словаря важной процедурой является проектирование базы данных, которое по мнению Е. Карпиловской [13], предусматривает два этапа:

1. Инфологический – этап сбора информации и ее структурирования, или этап внутренней формализации информации, моделирования содержания информации. Задача

этого этапа состоит в отборе объектов описания, типов информации об их строении и функционировании. Это этап изучения и описания определенной предметной области, ее внутренней формализации. Результатом работы на этом этапе является концептуальная информационная модель такой предметной области.

2. Датологический – этап оформления информации на соответствующем языке представления, пригодном для компьютерной обработки, или этап внешней формализации информации, моделирования ее формы, преобразования информации в данные. Задачей этого этапа является выработка способов представления объектов и информации о них в памяти компьютера, специальных маркеров-сигнализаторов для безошибочного «распознавания» компьютером того или иного типа информации, правил взаимодействия типов информации и получения из базы данных сведений в нужном виде или объеме, то есть это этап внешней формализации информации о языковых объектах.

Датологический этап, в свою очередь, содержит две стадии организации данных: логическую – адаптацию концептуальной информационной модели, полученной на инфологическом этапе, к операционной системе определенного типа компьютера и физическую – выбор рациональной структуры организации базы данных, методов работы с ней, исходя из возможностей, предоставляемых техническими параметрами.

Выводы. Итак, универсальной жесткой схемы этапов создания электронного словаря в современной литературе не существует. Учитывая опыт ученых и собственные исследования [15], предлагаем рассматривать такие условные этапы проектирования электронного словаря: аналитический, концептуальный, технологический, экспериментальный и рефлексивный. Характеристики предложенных этапов представлены в таблице 1.

Этап	Характеристика
аналитический	анализ реальной ситуации на рынке лексикографической продукции в конкретной области и реальных потребностей пользователей, определение актуальности и целесообразности создания нового словаря, исследование факторов, влияющих на выбор его параметров
концептуальный	разработка общей концепции словаря: определение его типа и назначения, установление принципов описания лексикографического материала, проектирование функций, макро- и микроструктуры словаря, отбор материала (первоисточники, печатные лексикографические работы, кодексы номенклатур, авторские наработки и т.д.).
технологический	программная реализация словаря, создание баз данных, конкретизация и детализация технических решений, разработка и дизайн интерфейса, налаживание работы программы, проверка функциональности программного обеспечения, создание руководства пользователя
экспериментальный	апробация и проверка работы бета-версии словаря
рефлексивный	корректировка, внесение необходимых изменений, устранение недостатков в функционировании электронного словаря, создание конечного продукта проектной деятельности, определение перспектив дальнейшего развития проекта

Такой подход отражает логику проектирования как процесса и предполагает изменение его состояний (этапы различаются по задачам, результатам и видам деятельности субъекта проектирования). Вместе с тем, подчеркиваем условный характер выделения этих частей, поскольку реальный процесс проектирования не имеет линейной последовательности прохождения этапов, находящихся в сложных отношениях взаимообусловленности и

взаимопроникновения. Разработка практически значимых моделей проектирования электронных словарей составляет перспективное направление дальнейших исследований.

Список литературы:

1. Палкова А. Основные понятия электронной лексикографии. Вестник ТвГУ, серия «Филология», 2015, № 4, 88-93.
2. Балалаева Е.Ю. Электронный словарь: сущность, структура, классификации. Современная педагогика, 2014, № 4. URL: <http://pedagogika.snauka.ru/2014/04/2238>.
3. Кудашев И. С. Проектирование переводческих словарей специальной лексики. Helsinki: Helsinki University Print, 2007.
4. Balalaieva O. Yu. Design of Electronic Textbooks in Latin Language for Higher Agricultural Educational Institutions. Ph.D. in Pedagogy. Kyiv, Institute of Information Technologies and Learning Tools of NAPS of Ukraine, 2016.
5. Табанакова В. Д., Ковязина М. А. Функциональная модель переводного специального словаря. Вестник ТюмГУ, 2006, № 4, 158-165.
6. Лінгвістично-інформаційні студії : Праці Українського мовно-інформаційного фонду НАН України : у 5 т. Т. 3 : Тлумачна лексикографія. Кн. 1 : Словник української мови у двадцяти томах / В. А. Широков та ін. Київ : Український мовно-інформаційний фонд НАН України, 2018.
7. Беляева Л. Н. Потенциал автоматизированной лексикографии и прикладная лингвистика. Известия РГПУ им. А.И. Герцена, 2010, № 134, 70-79.
8. Pastor V., Alcina A. Search techniques in electronic dictionaries: a classification for translators. International Journal of Lexicography. 2010. Vol. 23, No 3. P. 307-354.
9. Агапова Н.А., Картофелева Н.Ф. О принципах создания электронного словаря лингвокультурологического типа: к постановке проблемы. Вестник ТГУ, 2014, № 382, 6-10.
10. Шереметьева, С.О., Осминин, П.Г., Щербаков Е.С. К вопросу об электронных ресурсах профессиональной лексики. Вестник ЮУрГУ, серия: «Лингвистика», 2014, №1, 57-63.
11. Ефремова А.Н. Комплексный электронный словарь: особенности проектирования. Известия РГПУ им. А.И. Герцена, 2015, № 178, 39-43.
12. Загайнов С.С. Проект универсального электронного военного словаря. Томск: [б.и.], 2020. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vital:11608>.
13. Карпіловська Є.А. Вступ до комп'ютерної лінгвістики. Донецьк: ЮВ, 2006.
14. Major principles of teaching Latin at the Departments of Veterinary Medicine in Ukrainian agrarian universities. Edukacja – Technika – Informatyka. Wybrane problemy edukacji technicznej i zawodowej, 2013, No 4, Część 1, Rzeszów: Wyd. FOSZE, 529–534.
15. Balalaieva O. Design between Philosophy, Engineering and Pedagogy: A Systemic Approach. Euromentor, 2020, Vol. XI, No. 3, 89-104.

Article received: 2021-01-08