

Анна Филиппович

МАКЕТ ИНФОРМАЦИОННО-ПРОГРАММНОЙ СРЕДЫ ДЛЯ АВТОМАТИЗАЦИИ НАУЧНЫХ ИССЛЕДОВАНИЙ САР

Согласно представленному проекту НИР в Интегрированную ИПС для автоматизации исследований Словаря Академии Российской 1789–1794 гг. входят шесть компонент:

- 1) Лингвистическая база данных.
- 2) Программный комплекс автоматизированного ввода текста словарных статей в базу данных.
- 3) Программный комплекс создания электронных словарей, словоуказателей, словарных подмножеств, поисковых тезаурусов.
- 4) Двухуровневая информационно-поисковая система.
- 5) Биобиблиографическая информационная система «Создатель САР».
- 6) Гиперграфическая система факсимильной копии оригинального издания Словаря Академии Российской 1789-1794 гг.

Архитектура ИПС для автоматизации исследований САР включает в себя две части: первая — это программный комплекс для формирования электронного издания САР и его компонент, вторая — это само электронное издание (ЭСАР) Т.о. первая часть включает программные комплексы автоматизированного ввода текста словарных статей в базу данных и создания электронных словарей, словоуказателей, словарных подмножеств, поисковых тезаурусов. Они реализуются в виде программных компонент. Вторая часть содержит различные представления Словаря (в виде БД, факсимильной копии страниц, гипертекстовой системы и т.д.) и программный комплекс для работы с ними: для осуществления поиска, сортировки, фильтрации данных, для просмотра и формирование новых лингвистических объектов на основе данных САР.

Электронное издание Словаря Академии Российской включает в себя пять основных компонент (рис. 1):

- 1) гипертекстовую информационную систему (ГТИС САР),
- 2) гиперграфическую систему факсимильных копий страниц словаря (факсимильная копия страниц САР – ФК САР),
- 3) лингвистическую базу данных (ЛБД САР),

- 4) информационно-поисковую систему для исследования и обработки данных Словаря (ИПС САР),
- 5) биобиблиографическую информационную систему «Создатели САР» (БИС Создатели САР).

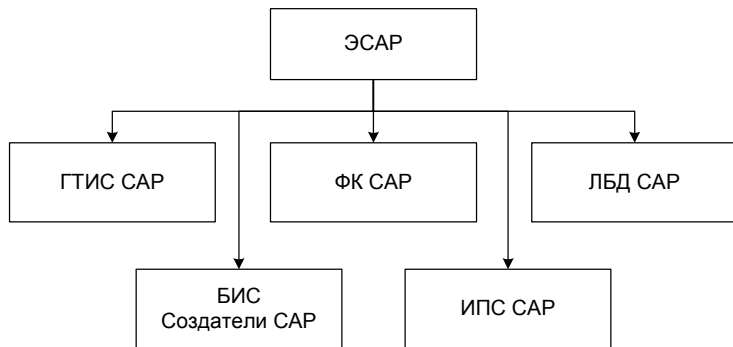


Рис. 1. Структурная схема ЭСАР.

Гипертекстовая информационная система

ГТИС САР «предназначена для простого, доступного для большинства непрофессиональных пользователей, доступа к содержанию САР» [2, с. 176]. Данная компонента должна выполнять следующие функции: вывод страниц гипертекстовой системы САР по томам и переход по разделам и по гиперссылкам в тексте.

Для формирования ГТИС САР используется электронная версия оригинал макета печатного переиздания словаря. Сам оригинал-макет сдавался в типографию в печатном виде. Переиздание САР содержит две части: *вступительную* с научным комментарием и *основную* с текстом словаря. Вначале — титульная часть словаря: фронтиспис — посвящение Е.Р. Дашковой, ее портрет; аван-титул; титульные листы; оборот титула. Далее рубрика «К читателю». После этого в печатном переиздании представлены научные статьи, посвященные словарю. Первая статья посвящена МГИ им. Дашковой. Далее идут несколько статей по тематике тома. Планируя переиздании, предполагалось, что каждый том будет посвящен какой-то теме. Завершает вступительную часть словаря рубрика «Российскому слову отличную пользу принесшему». Эта рубрика посвящена создателям словаря. Далее начинается сам словарь.

ГТИС САР должна повторять по содержанию печатное переизда-

ние. Далее представлен пример структуры разделов ГТИС 1-го тома САР:

- Титул;
- Оглавление;
- Оборот титула;
- К читателю;
- Московский гуманитарный институт им. Е.Р. Дашковой;
- О переиздании Словаря Академии Российской 1789-1794;
- Е.Р. Дашкова и словарь ее эпохи. XVIII век — взгляд из XX века;
- «Российскому слову отличную пользу принесемому»;
- Слово о Е.Р. Дашковой;
- Словарь
 - Титул;
 - Посвящение;
 - Предисловие;
 - Изъяснение;
 - Господа члены Российской Академии составляющие;
 - Показание;
 - Словарь А-В;
- Изменения в тексте переиздания;
- Список сокращений, использованных в тексте Словаря;
- Выходные данные.

ГТИС САР реализуется как компонента общей системы САР и может быть представлена как отдельная программа или как интернет версия.

Факсимильная копия

Факсимильная копия САР представляет собой набор изображений страниц Словаря Академии Российской. «Факсимиле — точное воспроизведение рукописи, документа, подписи, изображения» [3]. ЭСАР в первую очередь предназначен для проведения научных исследований и адресован ученым-лингвистам. Факсимильная копия позволяет увидеть оригинал словаря. Это является очень важным для проведения сопоставительных исследований. В процессе вычитки словаря было обнаружено множество ошибок, опечаток. Они были исправлены. Т.е. печатное переиздание полностью не соответствует оригиналу. «При переиздании первого тома было найдено компромиссное решение, основу которого составили предложения д.ф.н. М.И. Чернышевой. С одной стороны, сохранены особенности обозначений источников, оформления

заголовочных строк, отсылок и др. С другой стороны, сделаны следующие изменения: проведена унификация пунктуации в грамматических пометах и обозначениях источников — различные варианты использования разделителей (запятая, двоеточие, точка с запятой) заменены точкой; восстановлена пропущенная нумерация значений; откорректированы колонтитулы; исправлены явные опечатки (они, а также некоторые замеченные неточности и пояснения к ним сведены в таблицу изменений и представлены в конце тома). Кроме этого, решено было одновременно с корректурой проводить анализ опечаток и неточностей Словаря, а его результаты представить в других томах в форме научных комментариев или сопроводительных статей» [4, т. 1, с.413]. Кроме этого, нельзя исключить ошибки ввода и вычитки, сделанные при создании переиздание словаря. Т.о. включение компоненты ФК САР является необходимым.

Исходными данными для формирования факсимильной копии страниц САР являются страницы издания Словаря XVIII века и их ксерокопии, которые использовались для корректуры при создании печатного издания. Они должны быть отсканированы.

Анализ качества предварительно отсканированных страниц издания XVIII века и их ксерокопий показал, что требуется обработка изображений страниц, для этого может быть использован графический редактор Photoshop. Основными недостатками отсканированных изображений были следующие: слишком высокая или низкая контрастность, искажения (перекосы), наличие загрязнений и дефектов.

Для ликвидации этих недостатков следует разработать специальную технологию. Настройка контраста изображений и удаления основных загрязнений может осуществляться с помощью функции автоматизации в Photoshop — макросов. Остальные недостатки изображений — искажения, некоторые загрязнения и дефекты должны быть удалены вручную.

Далее необходимо разработать программную компоненту для просмотра и поиска страниц .

Т.о. основными задачами формирования гиперграфической информационной системы (ФК САР) являются:

1. Сканирование страниц САР, их ксерокопий.
2. Обработка изображений страниц.
3. Запись полученных изображений.
4. Разработка программной компоненты для просмотра и поиска страниц.

Библиографическая информационная система «Создатели САР»

Данная компонента содержит информацию о коллективе создателей САР. Исходными данными для формирования данной компоненты выступают материалы различных исследований жизни и деятельности персоналий.

В качестве вариантов для реализации библиографической информационной системы рассматриваются два варианта: в виде гипертекстовой системы, в виде базы данных. Основные функции системы: ввод и поиск информации о членах коллектива создателей САР.

Далее представлен вариант содержания описания каждого члена коллектива:

ФИО;

Портрет;

Факсимиле (подпись);

Звание;

Биография;

Работа, выполняемая для САР;

Основные труды;

Ссылки на дополнительную информацию.

Литература

1. *Филиппович Анна*. Электронная версия Словаря Академии Российской 1789-1794 годов. // Роль книгоиздания в развитии международных научных и культурных контактов: Материалы международной научной конференции (Москва, 21-23 сентября 2005 г.) / Сост. В.И.Васильев, М.А. Ермолаева, А.Ю. Самарин. — М.: Наука, 2005. — С. 293-296.

2. *Филиппович А.Ю.* Информационная технология создания электронного издания Словаря Академии Российской 1789-1794 гг. // Современные информационные технологии и письменное наследие: от древних рукописей к электронным текстам [Текст]: материалы междунар. науч. конф. (Ижевск, 13-17 июля 2006 г.) / отв. ред. В.А.Баранов. — Ижевск: Изд-во ИЖГТУ, 2006. — С. 174-178.

3. *Ожегов С.И. и Шведова Н.Ю.* Толковый словарь русского языка: 72500 слов и 7500 фразеол. выражений / Российская АН Ин-т рус.яз.; Российский фонд культуры. М.: Азъ Ltd., 1992. — 960 с.

4. *Филиппович Ю.Н.* Информационная технология переиздания «Словаря Академии Российской 1789-1794 гг.» // Книга и мировая

цивилизация: Материалы XI Междунар. Науч. Конф. По проблемам книговедения (Москва, 20-21 апр. 2004 г.): В 4 т. // [Сост. В.И. Васильев, М.А. Ермолаева, А.Ю. Самарин; Отв. Ред. В.И. Васильев, Б.В. Ленский]. — М.: Наука, 2004.