

Б. Б. Муминов, Д. Содикова

*Ташкентский Университет Информационных Технологий имени Мухаммада
ал-Хорезми, г. Ташкент*

УДК 004.4'2

ПРОЕКТИРОВАНИЕ БАЗЫ ДАННЫХ ПРИ ВНЕДРЕНИИ ТЕХНОЛОГИИ FSV

Аннотация. Разработано проектирование базы данных для внедрения технологии FSV в информационные системы и программные комплексы. Приведены структуры данных, которые должны быть в программном комплексе, и структуры данных, существующие в технологии FSV. Разработано на основе структуры данных, необходимых при интеграции, и их реляционных таблиц.

Ключевые слова. Проектирование, база данных, технология FSV.

Введение

В настоящее время задача поиска данных занимает почетное место в перечне актуальных прикладных задач. Для эффективного поиска данных, касающихся науки и образования, в система информационных ресурсов была создана технология FSV. Технология FSV – это инструментальная программная платформа, основанная на архитектуре клиент-сервер, с серверным приложением, интегрирующая и модифицирующая модели, способы и алгоритмы поиска и обработки данных в информационной среде. При этом F – способ представления запроса, то есть способ формирования выражения информационных потребностей пользователя системы. S – функция соответствия источника запросу, то есть степень соответствия (актуальности, совпадения) запроса и найденного источника. V – способ представления источников. Необходимо проектирование базы данных, основываясь на строгих правилах интегрирования технологии FSV с различными информационными системами и программным обеспечением.

Постановка задач: Проектирование базы данных (БД) с интеграцией технологии FSV с информационными системами и программными средствами.

Основная часть

Проектирование БД занимает особо важное место во внедрении технологии FSV в какой-либо программный комплекс. Потому что у технологии FSV имеется собственная БД, а программного комплекса, куда она должна быть внедрена, имеется собственная БД. При внедрении могут образоваться новые объединяемые, либо разделяемые структуры данных на основе требований и предложений пользователей системы. Для этого прежде всего требуется хорошее знание структур данных БД технологии FSV, а также БД программного комплекса, куда она должна быть внедрена, и модели IDEF1x (диаграмма ER). Поскольку заранее известно, что технология FSV будет внедряться в очень многие программные комплексы, в программном комплексе должны быть заданы множество основных источников и структуры данных для их полных текстов. Для общей картины, приведем данную структуру данных на основе таблицы (см. таблицы 1 и 2).

Таблица 1. Структура данных для множества источников

№	Наименование поля	Функции поля	Тип	Null	PK/ FK	Триггер
1	ID источника	Уникальный код источника	int	NN	PK	inc
2	Источник	Библиографическая квалификация источника	Nvarchar(1024)	NN		MARC21
3	TMID	Уникальный код полного текста источника	int	NN	FK	

Таблица 2. Структура данных для множества полного текста источников

№	Наименование поля	Функции поля	Тип	Null	PK/ FK	Триггер
1	IDTM	Уникальный код полного текста	int	NN	PK	Inc
2	fileName	Наименование полного текста	Nvarchar(250)	NN		
3	fullText	Полный текст	objectofTypeFile	NN		fileSystem

Для структуры данных, приведенной в таблицах 1 и 2, должна быть следующая модель IDEF1x.

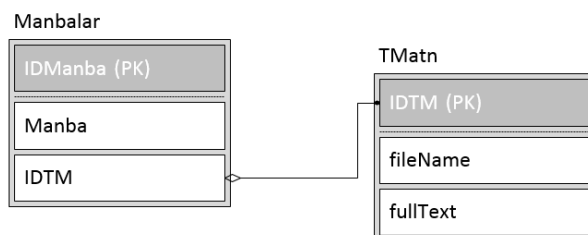


Рис. 1. Модель IDEF1x программного комплекса.

Модель IDEF1x приведенного программного комплекса и модели IDEF1x технологии FSV можно взаимно свести в несколько структур данных. Для этого необходимо знать основные структуры данных технологии FSV и их функции задачи (см. таблицу 3).

Таблица 3. Перечень структур данных технологии FSV.

№	Структура данных	Задачи и функции структуры данных
1.	Индекс терминов	Множества и индексы терминов и запросов
2	1-Семантические связи	Связи терминов, атрибутов и источников
3	Кл-во терминов	Кол-во терминов в источнике и значение g
4	2-Семантические связи	Семантические связи запроса и источника
5	Индекс непрозрачных терминов	Множество непрозрачных терминов и их связи
6	И-Семантическое ядро	Интеллектуальное семантическое ядро – на основе создания запросов и источников в НАМ
7	Временный индекс	Для создания индекса терминов при оптимизации запросов
8	Зоны и шаблоны	Соответствующие зоны запросов и множество их шаблонов
9	Логические действия	Логические действия и способы их записи
10	Источники	Перечень источников МИТ

Для сбора статических данных для интегрирования основных структур данных технологии FSV и структур данных программного комплекса необходимо создавать «Статические данные», для статической и динамической

отчетности «Сводные отчеты», для обмена и управления данными «Правила», для связывания структур данных «Новые связи», для временных данных «Временные» структуры данных. Каждая структура данных должна иметь собственный первоначальный ключ. Под первоначальным ключом понимается поле или группа полей, обеспечивающих уникальность информационного объекта, целесообразно, чтобы они состояли из количественных типов. Цель введения первоначального ключа в структуру данных заключается в обеспечении уникальности, упорядочения, сортировки и подбора при поиске данных.

Существует несколько способов использования структуры данных, то есть выполнения различных действий с данными в ней, они называются способами обращения к использованию данных. К способам обращения к использованию данных в технологии FSV мы можем отнести следующее:

1.Способ последовательных обращений. Способ последовательных обращений анализирует все записи в структуре данных для выполнения запросов к ней.

2. Способ индексных последовательных обращений. Способ индексных последовательных обращений направляет курсор записи к первой из записей, выполняющих условия запроса, для выполнения запроса, переданного в структуру данных. Затем курсор переходит к последующему ряду, удовлетворяющему условию запроса. Таким образом, определяются все ряды, удовлетворяющие условию запроса.

3.Способ прямого обращения. При способе индексных последовательных обращений из структуры данных отбираются напрямую соответствующие записи по значению одного поля или группы нескольких полей. Этот способ можно считать частным случаем второго способа. Потому что способ индексных последовательных обращений использует способ прямого обращения. Курсор направляется в первый ряд индекса, удовлетворяющего условию запроса, и после определения индексных записей,

удовлетворяющих необходимым условиям, передвигается по ряду индексов согласно способу последовательных обращений.

Приведем основные структуры данных, необходимые при внедрении технологии FSV в программный комплекс, на основе способов обращений (см. таблицы 4, 5, 6).

Таблица 4. Структура данных для сбора статических данных
«Статические данные»

Т.р	Наименование поля	Функции поля	Тип	Null	PK/ FK	Триггер
1	ID	Уникальный код	Int	NN	PK	Inc
2	staticParm	Произвольный параметр	Nvarchar(500)	NN		
3	Type	Какой именно параметр	Nvarchar(50)	NN		
4	staticID	Связи, если имеются	Int	N	FK	
5	isStatic	Для определения статичности или динамичности	Bin	NN		

Таблица 5. Структура данных для обмена и управления данными «Правила»

Т.р	Наименование поля	Функции поля	Тип	Null	PK/ FK	Триггер
1	ID	Уникальный код	Int	NN	PK	Inc
2	ruleIf	Правило	Nvarchar(50)	NN		
3	A	Действие для True	Navarchar(100)	NN		
4	B	Действие для False	Navarchar(100)	N		
5	Pram	Наличие параметра	Bin	NN		

Таблица 6. Структура данных для связывания структур данных «Новые связи»

Т.р	Наименование поля	Функции поля	Тип	Null	PK/FK	Триггер
1	ID	Уникальный код полного текста	Int	NN	PK	Inc
2	OneID	Первый объект	Int	NN	FK	
3	typeID1	Тип объекта	Nvarchar(100)	NN		
4	TwoID	Второй объект	Int	NN	FK	
5	typeID2	Тип объекта	Nvarchar(100)	NN		

Структура данных является мостом для обмена данными при внедрении технологии FSV в программный комплекс. В общем, для внедрения необходимо наличие в программном комплексе двух взаимосвязанных структур данных и трех вышеприведенных структур данных. Создание их необходимо для импорта структуры XML, созданной с помощью технологии FSV, а также требуется введение наименования и уникального поля одной основной структуры программного обеспечения.

Заключение

Внедрение технологии FSV в программных комплексах влечет за собой совершенствование модуля поиска данных. Это, наряду с удовлетворением потребностей пользователей программного комплекса, приведет к увеличению срока жизни (жизненный цикл) программного комплекса. После интегрирования интеллектуальное управление и создание семантических связей в составе технологии FSV, база непрозрачной информации, ядро базы информации начнет служить для программного комплекса.