

## Лабораторная работа №1/2 – представления и обработка данных

### Описание работы

Разработать способ организации данных в файле, позволяющий хранить, выбирать и гранулярно обновлять наборы записей общим объёмом от 10GB соответствующего варианту вида. Реализовать модуль или библиотеку для работы с ним в режиме курсора.

Используя данный способ сериализации, воспользоваться существующей библиотекой для описания схемы и реализации модуля, обеспечивающего функционирование протокола обмена запросами создания, выборки, модификации и удаления данных, и результатами их выполнения.

Использовать средство синтаксического анализа по выбору, реализовать модуль для разбора некоторого подмножества языка запросов по выбору в соответствии с вариантом формы данных. Должна быть обеспечена возможность описания команд создания, выборки, модификации и удаления данных.

Используя созданные модули разработать в виде консольного приложения две программы: клиентскую и серверную части. Серверная часть – получающая по сети запросы и операции описанного формата и выполняющая их над файлом, организованным в соответствии с разработанным способом. Имя фала данных для работы получать с аргументами командной строки, создавать новый в случае его отсутствия. Клиентская часть – получающая от пользователя команду, пересылающая её на сервер, получающая ответ и выводящая его в человекопонятном виде.

### Варианты заданий

Вариант	Форма данных	Протокол обмена
1	Документное дерево	Xml
2	Реляционные таблицы	Json
3	Граф узлов с атрибутами	Protocol Buffers
4	Документное дерево	Thrift
5	Реляционные таблицы	Xml
6	Граф узлов с атрибутами	Json
7	Документное дерево	Protocol Buffers
8	Реляционные таблицы	Thrift
9	Граф узлов с атрибутами	Xml
10	Документное дерево	Json
11	Реляционные таблицы	Protocol Buffers
12	Граф узлов с атрибутами	Thrift

## Расшифровка формы данных

	Документное дерево	Реляционные таблицы	Граф узлов с атрибутами
Организация элементов данных	Дерево узлов, несущих свойства	Таблицы записей, несущих поля	Граф узлов, несущих атрибуты
Природа связей	Физическая: родитель-ребенок	Логическая: по идентичным значениям	Физическая: связи между узлами
Примеры	Json, Xml, registry, прикладной уровень файловых систем	РСУБД, метаданные бинарных исполняемых модулей	Семантические сети, сетевые схемы
Состав схемы данных	Виды узлов, виды значений в узлах	Виды записей таблиц, виды значений в полях	Виды узлов, виды связей
Состав модели фильтра данных	Условия по содержанию элементов данных и отношениям между ними		
Примеры языков запросов	XPath, XQuery	SQL, LINQ	GraphQL, Cypher

## Описание примеров реализации

Организация страниц и данных в реестре Windows:

<https://github.com/msuhanov/regf/blob/master/Windows%20registry%20file%20format%20specification.md>

Организация данных в графовом хранилище Neo4j: <https://neo4j.com/developer/kb/understanding-data-on-disk/>

Организация страниц в MSSQL: <https://docs.microsoft.com/en-us/sql/relational-databases/pages-and-extents-architecture-guide?view=sql-server-2017>

Организация страниц в Postgre: <https://www.postgresql.org/docs/9.4/storage-page-layout.html>

Кучи данных в MSSQL: <https://docs.microsoft.com/en-us/sql/relational-databases/indexes/heap-tables-without-clustered-indexes?view=sql-server-2017#heap-structures>

Организация данных в Postgre: <http://rachbelaid.com/introduction-to-postgres-physical-storage/>, <http://www.interdb.jp/pg/pgsql01.html>

Организация данных в MongoDB: <https://www.quora.com/What-is-the-internal-file-structure-for-the-collection-in-MongoDB>, <https://docs.mongodb.com/manual/core/storage-engines/>, <http://bsonspec.org/spec.html>

Графы в MSSQL: <https://docs.microsoft.com/en-us/sql/relational-databases/graphs/sql-graph-architecture?view=sql-server-2017>