

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	6
1. Обзор предметной области	7
1.1. Биоинформатика	7
1.1.1. Анализ экспрессии генов	7
1.1.2. Используемые методы	7
1.2. Существующие решения для анализа экспрессии генов	8
1.2.1. R/Bioconductor	8
1.2.2. GENE-E	10
1.2.3. morpheus.js	10
1.2.4. ProjectX	13
1.3. Использование R в веб-разработке	14
1.3.1. Shiny	14
1.3.2. OpenCPU	14
1.4. Инфраструктура	16
1.4.1. Docker	16
1.4.2. Apache.....	17
1.5. Форматы сериализации данных	17
1.5.1. JSON	17
1.5.2. Protocol Buffers	18
1.6. Открытые источники данных экспрессии генов	18
1.6.1. Gene Expression Omnibus	18
1.6.2. The Cancer Genome Atlas	19
1.7. Постановка задачи.....	19
1.7.1. Цель работы	19
1.7.2. Основные задачи.....	19
1.7.3. Требования к веб-приложению phantusus.....	19
Выводы по главе 1.....	20
2. Архитектура проекта phantusus.....	21
2.1. Общая схема	21
2.2. Взаимодействие между клиентом и сервером.....	21
2.2.1. OpenCPU	22
2.2.2. Protocol Buffers	23

2.3. Поддержка ExpressionSet	24
2.3.1. Поддержка ExpressionSet на стороне клиента	24
2.3.2. Создание ExpressionSet из внешних данных на стороне сервера	25
2.4. Способы визуализации	26
2.4.1. Отрисовка графиков на стороне сервера — статические изображения	26
2.4.2. Интерактивные графики — plotly.js	26
Выводы по главе 2	26
3. Реализация и использование	27
3.1. Реализованные методы анализа экспрессии	27
3.1.1. Метод главных компонент и визуализация его результата	27
3.1.2. Кластеризация методом kmeans	28
3.1.3. Анализ дифференциальной экспрессии	28
3.1.4. Дифференциальная экспрессия — limmaAnalysis	29
3.1.5. Загрузка данных из GEO — loadGEO	30
3.2. Структура git-репозитория	32
3.3. Кэш для данных из GEO	32
3.4. Единый R-пакет phantasus	33
3.5. Docker-образ phantasus	33
3.5.1. Запуск Docker-контейнера	34
3.6. Настройка с помощью Apache	34
3.6.1. Переадресация OpenCPU-сервера	34
3.6.2. Балансировщик для multi-user соединения	35
3.7. Статистика использования	36
Выводы по главе 3	36
ЗАКЛЮЧЕНИЕ	37
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	38
ПРИЛОЖЕНИЕ А. Протокол сериализации в ProtoBuf	41
ПРИЛОЖЕНИЕ Б. Dockerfile	42