

Санкт-Петербургский национальный исследовательский университет  
информационных технологий, механики и оптики  
Факультет информационных технологий и программирования  
Кафедра компьютерных технологий

**Реализация эффективного взаимодействия между  
платформой для анализа экспрессии генов Morpheus и  
библиотекой вычислительных методов R/Bioconductor**

Зенкова Д.М.

Научный руководитель: Сергушичев А. А.

Санкт-Петербург  
2017



# ОГЛАВЛЕНИЕ

Стр.

<b>ВВЕДЕНИЕ .....</b>	<b>5</b>
<b>ГЛАВА 1. ОБЗОР ПРЕДМЕТНОЙ ОБЛАСТИ</b>	<b>6</b>
1.1. Биоинформатика .....	6
1.1.1. Анализ экспрессии генов.....	6
1.1.2. Используемые методы .....	6
1.2. Существующие решения для анализа экспрессии генов ...	6
1.2.1. GENE-E .....	6
1.2.2. morpheus.js .....	6
1.2.3. R/Bioconductor.....	7
1.3. Инструменты, которые могут быть применены .....	7
1.3.1. Язык R и библиотека Bioconductor.....	7
1.3.2. JavaScript .....	7
1.3.3. R shiny .....	7
1.3.4. OpenCPU .....	7
1.3.5. Gene Expression Omnibus.....	7
1.3.6. Docker .....	8
1.3.7. Protocol Buffers .....	8
1.3.8. Apache2.....	8
1.3.9. HTML.....	8
1.4. Постановка задачи .....	8
Выводы по главе 1.....	8
<b>ГЛАВА 2. АРХИТЕКТУРА ПРОЕКТА</b>	<b>9</b>
Резюме .....	9
<b>ГЛАВА 3. РЕАЛИЗАЦИЯ</b>	<b>10</b>
Резюме .....	10
<b>ЗАКЛЮЧЕНИЕ .....</b>	<b>11</b>
<b>СПИСОК ИСТОЧНИКОВ .....</b>	<b>12</b>

## **ВВЕДЕНИЕ**

## ГЛАВА 1. ОБЗОР ПРЕДМЕТНОЙ ОБЛАСТИ

### 1.1. Биоинформатика

#### 1.1.1. Анализ экспрессии генов

#### 1.1.2. Используемые методы

### 1.2. Существующие решения для анализа экспрессии генов

#### 1.2.1. GENE-E

Платформа для анализа данных и визуального исследования данных, созданная на Java и R. Содержит в себе множество полезных для исследования инструментов: тепловые карты, кластеризацию, фильтрацию, построение графиков и т.д. Позволяет исследовать любые данные в виде матрицы. К тому же, содержит дополнительные инструменты для геномных данных.

Недостатки:

- Чтобы использовать, необходимо устанавливать на свой компьютер.
- Поддержка данного приложения прекратилась в связи с созданием morpheus.js
- Не имеет открытого исходного кода, а только API.

#### 1.2.2. morpheus.js

Веб-приложение от создателя GENE-E. Создано с теми же целями и мотивацией, но уже на JavaScript и с открытым исходным кодом. Удобно для использования исследователями без навыков программирования и так же, как и GENE-E, применимо к любым матрицам.

Недостатки:

- Ограниченный набор функций, которых недостаточно для полноценного анализа.
- Для расширения биоинформатическими алгоритмами требуется реализовывать их заново на JavaScript.

### **1.2.3. R/Bioconductor**

R - язык программирования для статистического анализа данных и работы с графикой. Bioconductor - библиотека, содержащая в себе множество реализаций биоинформатических алгоритмов и методов обработки биологических данных на R. Она постоянно обновляется, пополняется новыми библиотеками, модерируется сообществом. R и Bioconductor очень популярны в биоинформатической среде ввиду предоставляемых возможностей.

Однако для качественного и полноценного анализа с помощью этих инструментов, нужно иметь навыки программирования на R, что весьма неудобно для исследователей биологических специальностей.

## **1.3. Инструменты, которые могут быть применены**

### **1.3.1. Язык R и библиотека Bioconductor**

### **1.3.2. JavaScript**

JavaScript - язык программирования, широко используемый для написания веб-приложений.

### **1.3.3. R shiny**

### **1.3.4. OpenCPU**

OpenCPU - система для встроенных научных вычислений и воспроизводимых исследований, предоставляющая HTTP API для взаимодействия с R-серверами. Имеется также библиотека `орепсри.js` для интеграции JavaScript и R.

### **1.3.5. Gene Expression Omnibus**

GEO - публичный репозиторий с геномными данными.

В библиотеке Bioconductor есть R-пакет `GEOquery` для удобной загрузки данных из GEO.

**1.3.6. Docker**

**1.3.7. Protocol Buffers**

**1.3.8. Apache2**

**1.3.9. HTML**

**1.4. Постановка задачи**

**Выводы по главе 1**

## **ГЛАВА 2.   АРХИТЕКТУРА ПРОЕКТА**

В этой главе будут подробно рассмотрены элементы проекта, их взаимосвязь и ключевые для архитектуры выдержки из исходного кода.

### **Резюме**



## ГЛАВА 3. РЕАЛИЗАЦИЯ

### Резюме

## **ЗАКЛЮЧЕНИЕ**

## **СПИСОК ИСТОЧНИКОВ**