**Спецификация требований к программному обеспечению**

Информация о документе:  
  
Идентификатор документа: SRS-2023-HH  
Редакция: 1.0  
Дата: 2023-02-06  
Автор: Алексей Боровой

[1. Введение 3](#_Toc398378400)

[Цели проекта 3](#_Toc1839241101)

[Соглашения о терминах 3](#_Toc1245569589)

[Предполагаемая аудитория и последовательность восприятия 3](#_Toc617491367)

[Масштаб проекта 3](#_Toc1536422865)

[Ссылки на источники 3](#_Toc1315720527)

[2. Общее описание 4](#_Toc675784130)

[Видение продукта 4](#_Toc2010347574)

[Функциональность продукта 5](#_Toc1551457967)

[Функциональные требования: 5](#_Toc607609495)

[Классы и характеристики пользователей 5](#_Toc808040980)

[Среда функционирования продукта (операционная среда) 5](#_Toc58167232)

[Рамки, ограничения, правила и стандарты 5](#_Toc511608839)

[Документация для пользователей 5](#_Toc85234485)

[Допущения и зависимости 5](#_Toc164396846)

[3. Функциональность системы 5](#_Toc1353820680)

[Функциональный блок 1. Сбор данных 5](#_Toc1074928795)

[Функциональный блок 2. Оперативное (промежуточное) хранение данных 6](#_Toc1695819001)

[Функциональный блок 3. Очистка и предобработка информации для помещения в базу данных 6](#_Toc226986823)

[Функциональный блок 4. Создание базы данных 6](#_Toc1036577089)

[Функциональный блок 5. Помещение обработанной информации в базу данных 6](#_Toc1363790359)

[Функциональный блок 6. Создание интерактивного дашборда. 6](#_Toc1131599282)

[Функциональный блок 7. Контейнеризация и запуск приложения. 6](#_Toc673954742)

[Функциональный блок 8. Отладка. Поиск и устранение ошибок и неисправностей в работе приложения. 6](#_Toc1328594516)

[4. Требования к внешним интерфейсам 7](#_Toc153068906)

[5. Нефункциональные требования 7](#_Toc2091177683)

[6. Прочее 7](#_Toc705951379)

# Введение

В этом документе представлена полная спецификация требований к программному обеспечению для проекта, направленного на разработку приложения для сбора, обработки, хранения и визуализации информации о вакансиях по специальностям “Аналитик” (Data Analyst) (все направления), “Инженер данных” (Data engineer), “Дата саентист” (Data Dcientist).

## Цели проекта

* Получение исчерпывающей и наглядной информации о распределении вакансий по направлениям аналитики данных (в количественном и процентном выражении).
* Определение для каждого направления аналитика данных ключевых навыков, подсчёт количественного выражения упоминаний тех или иных навыков в вакансиях, составление рейтинга таких навыков по частоте упоминания.
* Определение распределения вакансий по: географии, отраслям бизнеса, уровню подготовки специалистов (грейд).

## **Соглашения о терминах**

Нужно ли?

## Предполагаемая аудитория и последовательность восприятия

Предполагаемой аудиторией (конечными пользователями) приложения являются соискатели на вакантные должности по аналитическим направлениям.

## Масштаб проекта

В данном проекте задействована рабочая группа, состоящая из 5 (пяти) сотрудников: лидера команды (Team Lead), системного аналитика, двух дата инженеров и аналитика данных. Бюджет проекта - 0 руб. Сроки проекта - 3 месяца. Границы проекта: от сайта-агрегатора вакансий (hh.ru) до конечного потребителя визуальной информации на интерактивном дашборде.

## Ссылки на источники

https://github.com/mgarkunov/hh\_ru\_analytics

# **Общее описание**

## Видение продукта

В настоящее время на сайтах-агрегаторах каждый день размещается множество вакансий по различным направлениям бизнеса, промышленности, торговли, услуг и проч. В виду того, что инструментов, которыми должен обладать соискатель на желаемую должность весьма большое количество, такой соискатель может испытывать трудности с определении перечня необходимых к изучения технологий. Данный проект позволит получить распределение по необходимым инструментам и технологиям в зависимости от искомой должности и выделит наиболее популярные у работодателей, что позволит соискателю чётко определить для себя приоритет в изучении технологий.

Кроме того, в данном проекте проводится аналитика распределения технологий в зависимости от географического расположения работодателя, что позволит сегментировать технологии по населенным пунктам.

В проекте так же предусматривается систематическая выгрузка данных с сайта-агрегатора, очистка и аккумулирование информации в базе данных, после чего данные размещаются в интерактивном обновляемом дашборде.

## Функциональность продукта

## Функциональные требования:

## Классы и характеристики пользователей

## Среда функционирования продукта (операционная среда)

## Рамки, ограничения, правила и стандарты

## Документация для пользователей

## Допущения и зависимости

# **Функциональность системы**

## Функциональный блок 1. Сбор данных

* Описание и приоритет.

Данный блок представляет собой сбор данных с сайта <https://hh.ru/> через предлагаемый сайтом API с помощью запросов к сайту в среде Python 3.

Приоритет - высокий.

* Функциональные требования:

1. Выполнен запрос к сайту через API сайта hh.ru.
2. Получен ответ от сайта (response code 200).
3. Скрипт контролирует количество неуспешных попыток авторизации в рамках одной сессии. В случае превышения лимита попыток авторизации сессия закрывается. Лимит безуспешных попыток авторизации (получения ответа) - 10 с промежутком 15 секунд.
4. В случае запрета сервера на запросы клиента по причине достижения лимита авторизаций, используется proxy-маска.
5. Получены данные в формате XML, содержащие [УТОЧНИТЬ ЧТО БУДЕМ ЗАПРАШИВАТЬ].
6. К каждой строке (запросу) формируется токен, состоящий из даты и времени формирования запроса. Формат: ‘yyyy:mm:dd hh:mm:ss’.

## Функциональный блок 2. Оперативное (промежуточное) хранение данных

* Описание и приоритет.

Данный блок представляет собой алгоритм сохранения данных промежуточного хранения полученных через API данных с сайта hh.ru.

Приоритет - высокий.

* Функциональные требования:

1. На локальном сервере (локальный персональный компьютер) создан файл промежуточного хранения. Формат .txt. Название: ‘temporary\_db.txt’.
2. Данные по каждому запросу формируются в ответ в формате XML, после чего производится распарсинг в тип “строка”.
3. Строка конкатенируется со сгенерированным токеном.
4. Сформированная структура данных копируется в ‘temporary\_db.txt’, файл сохраняется.

## Функциональный блок 3. Очистка и предобработка информации для помещения в базу данных

## Функциональный блок 4. Создание базы данных

## Функциональный блок 5. Помещение обработанной информации в базу данных

## Функциональный блок 6. Создание интерактивного дашборда.

## Функциональный блок 7. Контейнеризация и запуск приложения.

## Функциональный блок 8. Отладка. Поиск и устранение ошибок и неисправностей в работе приложения.

# **Требования к внешним интерфейсам**

**Интерфейсы пользователя (UX)**

**Программные интерфейсы**

**Интерфейсы оборудования**

**Интерфейсы связи и коммуникации**

# **Нефункциональные требования**

**Требования к производительности**

**Требования к сохранности (данных)**

**Требования к качеству программного обеспечения**

**Требования к безопасности системы**

**Требования на интеллектуальную собственность**

# **Прочее**

**Приложение А: Глоссарий**

**Приложение Б: Модели процессов и предметной области и другие диаграммы**

**Приложение В: Список ключевых задач**