

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ РАДИОФИЗИКИ И
КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ

Кафедра системного анализа и компьютерного моделирования

**РАЗРАБОТКА ВЕБ-ПЛАТФОРМЫ ДЛЯ ХРАНЕНИЯ И ОБРАБОТКИ
БИОИНФОРМАЦИОННЫХ ДАННЫХ**

Аннотация к магистерской диссертации

специальность «Радиофизика»

Павлова Романа Павловича
студента 2 курса,
специальность
«радиофизика»
Научный руководитель:
Гилевский С.В.

Минск – 2021

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Данная работа посвящена разработке веб-платформы для эффективного и безопасного решения задач биоинформатики, хранения и обработки исходных данных, ориентируясь на внутреннее использование в учреждениях здравоохранения и возможность модульного расширения системы.

В ходе работы произведены анализ требований к функционалу и информационной безопасности разрабатываемой веб-платформы, обзор существующих решений; проектирование и прототипирование; апробация разработанного прототипа, анализ полученных результатов.

Результаты имеют практическую ценность для применения в учреждениях здравоохранения, производящих анализ данных, полученных в ходе секвенирования генома.

Для разработки клиентской логики системы использован язык JavaScript, для создания пользовательского интерфейса были применены JavaScript фреймворк Vue.js, ориентированный на архитектуру MVVM и тулkit Bootstrap.

Серверная часть имеет микросервисную архитектуру и была разработана Лабутиным Михаилом Дмитриевичем на языке Java с использованием фреймворка Spring Boot.

ИСПОЛЬЗУЕМЫЕ АББРЕВИАТУРЫ

MVVM — Model - View - ViewModel

NGS — Next Generation Sequencing

ОС — Операционная Система

API — Application Programming Interface

JWT — Java Web Token

TOTP — Time-based One-Time Password

РЕФЕРАТ

Диссертация: 52 страницы, 25 иллюстраций, 19 источников, 3 приложения.

Ключевые слова: биоинформатика, анализ данных, веб-платформа.

Объект исследования - обработка биоинформационных данных.

Цель работы – разработка веб-платформы для хранения и обработки биоинформационных данных.

В ходе деятельности современных онкодиспансеров неприменно возникает необходимость распознавания природы опухолей и изучения их развития. Путем секвенирования генома получают исходные данные об опухолевых клетках, в дальнейшем полученная информация подвергается различной обработке. Выполнение алгоритмов биоинформатики требует значительное количество вычислительных операций и операций доступа к хранилищу, в связи с чем запуск данных алгоритмов на персональных рабочих станциях представляется нецелесообразным, гораздо эффективнее обрабатывать результаты секвенирования генома человека на выделенном высокопроизводительном сервере.

В целях упрощения доступа к централизованному серверу обработки биоинформационных данных было решено разработать веб-платформу, позволяющую всем специалистам онкологического центра использовать её для запуска обработки, просмотра его статуса и результатов. Поскольку применять данную систему планируется в медицинских учреждениях для обработки личных медицинских данных пациентов, остро встает необходимость обеспечения безопасности системы, защиты ее и обрабатываемых данных от несанкционированного доступа. Пристальное внимание уделено вопросам аутентификации пользователя и связи веб-интерфейса и сервера. Система разрабатывалась в сотрудничестве с заведующим лаборатории молекулярно-генетических исследований РНПЦ детской онкологии, гематологии и иммунологии Мигасом Александром Александровичем в целях ее дальнейшего практического применения.

РЭФЕРАТ

Дысертация: 52 старонкі, 25 ілюстрацый, 19 крыніц, 3 дадатку.

Ключавыя слова: біяінфарматыкі, аналіз дадзеных, вэб-платформа.

Аб'ект даследавання - апрацоўка биоинформационных дадзеных.

Мэта працы - распрацоўка вэб-платформы для захоўвання і апрацоўкі біяінфармацыйных дадзеных.

У ходзе дзейнасці сучасных анкарадыспансераў недастасоўна ўзнікае неабходнасць распознання прыроды пухлін і вывучэння іх развіцця. Шляхам секвеніраваннягеному атрымліваюць зыходныя дадзеные аб опухолевых клетках, у далейшым атрыманая інфармацыя падвяргаецца рознай апрацоўцы. Выкананне алгарытмаў біяінфарматыкі патрабуе значнай колькасці вылічальных аперацый і аперацый доступу да сховішча, у сувязі з чым запуск дадзеных алгарытмаў на персанальных працоўных станцыях ўяўляеца немэтазгодным, значна больш эфектыўна апрацоўваць вынікі секвеніравання геному чалавека на вылучаным высокапрадуктыўным серверы.

У мэтах спрашчэння доступу да цэнтралізаванага сервера апрацоўкі біяінфармацыйных дадзеных было вырашана распрацаваць вэб-платформу, якая дазваляе ўсім спецыялістам анкалагічнага цэнтра выкарыстоўваць яе для запуску апрацоўкі, прагляду яго статусу і вынікаў.

Паколькі ўжываць дадзеную сістэму плануеца ў медыцынскіх установах для апрацоўкі асабістых медыцынскіх дадзеных пацыентаў, востра паўстае неабходнасць забеспячэння бяспекі сістэмы, абароны яе і апрацоўваных дадзеных ад несанкцыянаванага доступу. Пільная ўвага нададзена пытанням аўтэнтыфікацыі карыстальніка і сувязі вэб-інтэрфейсу і сервера. Сістэма распрацоўвалася ў супрацоўніцтве з загадчыкам лабараторыі малекулярна-генетычных даследаванняў РНПЦ дзіцячай анкалогіі, гематалогіі і імуналогіі Мігасам Аляксандрам Аляксандравічам у мэтах яе далейшага практычнага прымяняння.

ABSTRACT

Dissertation: 52 pages, 25 illustrations, 19 sources, 3 appendices.

Keywords: bioinformatics, data analysis, web platform.

The object of research is the processing of bioinformatic data.

The purpose of the work is to develop a web platform for storing and processing bioinformatic data.

In the course of the activity of modern oncologic dispensaries, it is inevitable that it becomes necessary to recognize the nature of tumors and study their development. By sequencing the genome, the initial data on the tumor cells is obtained, and then the information obtained is subjected to various processing. The implementation of bioinformatics algorithms requires a significant number of computational operations and storage access operations, which makes it impractical to run these algorithms on personal workstations, and it is much more efficient to process the results of human genome sequencing on a dedicated high-performance server.

In order to simplify access to a centralized bioinformatic data processing server, it was decided to develop a web platform that allows all specialists of the oncology center to use it to start processing, view its status and results. Since it is planned to use this system in medical institutions for processing personal medical data of patients, there is an urgent need to ensure the security of the system, protect it and the processed data from unauthorized access. Close attention is paid to the issues of user authentication and the connection between the web interface and the server.

The system was developed in cooperation with the head of the Laboratory of Molecular and Genetic Research of the Republican Scientific and Practical Center of Pediatric Oncology, Hematology and Immunology, Migas Alexander Alexandrovich, for its further practical application.