МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**(ФГБОУ ВО «КубГУ»)**

**Факультет компьютерных технологий и прикладной математики**

**Кафедра вычислительных технологий**

**КУРСОВАЯ РАБОТА**

**РАЗРАБОТКА ПРОГРАММЫ ДЛЯ ПОИСКА СОТРУДНИКОВ НА САЙТАХ ВАКАНСИЙ**

Работу выполнил \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Л.Т. Алексанянц (подпись, дата)

Направление подготовки 02.03.02 «Фундаментальная информатика и информационные технологии» курс 3

Направленность (профиль) «Вычислительные технологии»

Научный руководитель

канд. техн. наук доц.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Т.А.Приходько (подпись)

Нормоконкролер

ассистент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_А.А. Климец (подпись)

Краснодар 2019

СОДЕРЖАНИЕ

[ВВЕДЕНИЕ 3](#_Toc26957236)

[1. Обзор рынка на наличие готовых решений 4](#_Toc26957237)

[1.1 Обзор готовых решений для автоматизированного поиска сотрудников 4](#_Toc26957238)

[1.1.1 E-Staff 4](#_Toc26957239)

[1.1.2 Experium 4](#_Toc26957240)

[1.1.3 Резюмакс 5](#_Toc26957241)

[1.1.4 Staffery 2009 5](#_Toc26957242)

[1.1.5 Oracle Taleo Cloud Service 6](#_Toc26957243)

[1.1.6 Microsoft Dynamic CRM Кадровое агентство 6](#_Toc26957244)

[1.1.7 1С: Предприятие 8. Кадровое агентство 6](#_Toc26957245)

[1.1.8 Molga (SAP) 7](#_Toc26957246)

[1.2 Анализ рынка ПО по автоматизированному сбору резюме 7](#_Toc26957247)

[1.3 Анализ функциональности представленных решений 8](#_Toc26957248)

[1.3.1 Полезный функционал для поиска сотрудников 8](#_Toc26957249)

[1.3.2 Вспомогательный функционал для решения кадровых проблем компании 8](#_Toc26957250)

[2. Стек технологий для разработки и внедрения автоматизированного поиска резюме на различных работных сайтах по заданным параметрам 11](#_Toc26957251)

[2.1 Технологии Web Scrapping 11](#_Toc26957252)

[2.2 Проблемы автоматизированного поиска 13](#_Toc26957253)

[2.3 Обзор библиотеки BeautifulSoup. 13](#_Toc26957254)

[3. Инструменты диагностики программ для WEB 15](#_Toc26957255)

[3.1 Анализ запроса с использованием функционала браузера Mozilla Firefox 15](#_Toc26957256)

[3.2 Анализ запросов с использованием программы HTTP Analyzer v7 17](#_Toc26957257)

[4. Программа для поиска резюме 19](#_Toc26957258)

[4.1 Поиск резюме на сайте domkadrov.ru 19](#_Toc26957259)

[4.2 Поиск резюме на сайте avito.ru 21](#_Toc26957260)

[ЗАКЛЮЧЕНИЕ 23](#_Toc26957261)

[ЛИТЕРАТУРНЫЕ ИСТОЧНИКИ 24](#_Toc26957262)

## ВВЕДЕНИЕ

В настоящее время существует множество компаний, которые растут и развиваются. С увеличением масштабов компании растёт и количество сотрудников, работающих в ней. Несомненно, поиск сотрудников – важная часть любой компании и чем быстрее организация привлечет к себе необходимые кадры, тем стремительнее будет её рост.

Не для кого не секрет, что в крупных компаниях существуют специальные отделы по поиску сотрудников именуемые «отдел кадров». Чем больше компания, тем больше времени приходится тратить на поиски сотрудников. Отдел со временем перестаёт справляться с нагрузкой и приходится расширять его, а значит необходимо платить еще больше зарплат людям, занимающимся поиском других людей.

Возникает вопрос: возможно ли автоматизировать поиск сотрудников? Решение данной проблемы позволит сократить количество зарплат в «отделе кадров» и увеличить скорость, а главное плодотворность поиска сотрудников. Сэкономленные деньги можно будет потратить на разработку и поддержку продукта.

Курсовая работа выполняется в рамках кейса с сайта профстажировки.рф. Работа выполнялась с учетом требований создателя кейса. Так же были произведены дополнительные попытки к дальнейшему возможному развитию проекта. Развитие проекта напрямую зависит от желания создателя кейса.

## Обзор рынка на наличие готовых решений

На сегодняшний день уже существует множество готовых решений по автоматизированному поиску сотрудников. Каждое имеет свои достоинства и недостатки.

В рамках данной работы были рассмотрены самые популярные из них.

## 1.1 Обзор готовых решений для автоматизированного поиска сотрудников

## 1.1.1 E-Staff

Это одно из старейших и самых популярных ПО на рынке. Софт можно использовать и как комплексную программу для сотрудника HR-отдела, и как универсальный инструмент для крупных кадровых агентств. Интернет-модуль имитирует действия человека, экономя время и усилия на заполнение форм и обработку результатов, выдаваемых сайтами. В число опций входит автоматическая публикация вакансий на сайтах по поиску работы (в основном списке их более 40) и сбор откликов от соискателей. Программа ведет учет клиентов, вакансий и заявок. Удобный поиск резюме по заданным критериям. Ведется статистика по работе рекрутерской компании и по занятости сотрудников. Разумеется, есть функция импорта резюме кандидатов из почтовых сервисов и документов Word или OpenOffice и интеграция с другими системами (Босс-Кадровик, 1С, SAP, WebSoft, WebTutor).

Цена программы для компании начинается от 150 тысяч рублей.

## 1.1.2 Experium

Сервисные возможности программы позволяют публиковать и анализировать вакансии на популярных сайтах, собирать отклики прямо из программы, создавать и вести базу данных не только кандидатов, но и собственных специалистов компании. Это не только софт для рекрутинга, но и полноценная программа для управления внутренними ресурсами: здесь можно хранить данные об организационной структуре и сотрудниках компании, работать с массивами документов в текстовых и графических форматах, формировать отчетность по любым параметрам. Ежедневник Experium синхронизируется с Outlook, Mozilla или Google-календарем.   
 Цена программы варьируется от 200 000 рублей до 1200 000 рублей.

## 1.1.3 Резюмакс

Онлайн-сервис для HR-менеджеров и соискателей. На первый взгляд, эта программа едва перескочила кустарный уровень развития. На деле софт может приятно удивить очевидной простотой и функциональностью. Предусматривает создание на сайте компании раздела, который способен автоматически размещать вакансии, принимать анкеты соискателей и проводить предварительное тестирование. Базовые опции: уведомление о поступлении новых кандидатов, размещении новых заявок или проведении собеседований. Автоматический ответ по электронной почте, вполне вменяемый поиск данных по заданным критериям и ключевым словам. А главное преимущество — это облачный сервис, где нет привязки к конкретным машинам: работайте из дома, в дороге, из другого офиса. Также можно отметить мультиязычный интерфейс, поддержку HR-XML, возможность настройки на несколько системных уровней, когда каждый пользователь имеет индивидуальный доступ к тем функциям, которые ему предоставлены.   
 Цена программы: от 50 000 руб. (или 3000 руб./мес.)

## 1.1.4 Staffery 2009

Это приложение для управления данными и отслеживания информации об образовании, профессиональном опыте и навыках сотрудников. Программа значительно упрощает работу HR-специалистам компании. Помимо хранения персональных данных облегчает быстрый поиск сотрудников по выбранным вручную критериям. Также позволяет сотрудникам самим вносить изменения в свою персональную карточку или резюме. Приложение сертифицировано корпорацией Microsoft в ходе тестирования партнерских продуктов.  
 Цена программы: от 8499 руб.

## 1.1.5 Oracle Taleo Cloud Service

Одна из самых комплексных, продвинутых и новейших облачных платформ в HR-сфере. Софт способен накапливать всю информацию о сотрудниках на протяжении их трудовой биографии, очень приличная навигация по кадровому потенциалу компании. На основании собираемых данных программа позволяет прогнозировать возможные потребности в специалистах различных отраслей (хотя прогноз получается приблизительным). Это Oracle, так что интерактивная структура приложения предназначена для компаний самого широкого масштаба и из любых отраслей. Taleo — один из мировых лидеров по числу пользователей.  
 Цена программы на рынке от 7000 рублей.

## 1.1.6 Microsoft Dynamic CRM Кадровое агентство

Это решение типовых отраслевых задач в сфере рекрутинга: подбор сотрудников, автоматизация взаимодействия работодателя и соискателей, контроль сроков исполнения контрактов, контроль заказов руководителей компаний и др. К очевидным плюсам можно отнести ведение общей базы данных, касающейся обращений, договоров, вакансий, использования автоматизированного подбора соискателей и кандидатов по конкретным параметрам. Для HR-менеджеров настроена автоматизация их повседневной работы, удобная навигация, быстрый доступ к информации о наличии свободных вакансий, мониторинг выполнения задач и т д.  
 Цена одной лицензии от 220 тысяч рублей**.**

## 1.1.7 1С: Предприятие 8. Кадровое агентство

Основная функция модуля — автоматизация повседневной работы в рекрутинговых и кадровых агентствах. Максимально упрощены процессы оценки профессиональных знаний, умений, индивидуальных качеств кандидатов. При разработке ПО ставилась задача облегчить бизнес-процессы рекрутинговых агентств и кадровых служб, автоматизировать взаимоотношения между клиентами и компанией. Не самый удачный с точки зрения юзабилити, но один из самых адаптированных под российские условия сайт. В число сильных сторон входит ведение общей информационной базы кандидатов на вакансии, автоматическое сопоставление требований работодателей и возможностей соискателей. Целевой аудитории — HR-менеджерам — софт предлагает удобную обработку информации, повышение оперативности в управлении работой и снижение вероятности ошибок при работе с клиентами.

Цена модуля начинается от 27 000 руб. на одного пользователя.

## 1.1.8 Molga (SAP)

Молодая разработка известной международной компании SAP AG.. Малый и средний бизнес от покупки решений этой компании, вероятно, откажется из-за неподъемного бюджета. Приложение, помимо поиска и подбора сотрудников,  включает в себя разработку оптимальной модели управления персоналом, унификацию процессов расчета заработной платы, учет стратегии предприятия, касающейся вопросов HR и многое другое.  
 Цена программы: от 3 000 000 руб. (базовая функциональность)

## 1.2 Анализ рынка ПО по автоматизированному сбору резюме

Анализ вышеперечисленных решений показывает, что покупка программы для автоматизации подбора персонала обойдется компании как минимум в 10 000 рублей. Нужно учесть тот факт, что большинство решений уже не поддерживаются разработчиками и в случае возникновения ошибок рассчитывать на квалифицированную помощь не приходится. Стоит отметить, что, покупая готовое решение, компания теряет в «гибкости», ведь если надо будет добавить либо изменить готовый функционал, далеко не каждый разработчик готов на это пойти. Было замечено, что львиную долю компания будет переплачивать за функционал, которым она даже не будет пользоваться.

Исходя из этого были подведены итоги: разработать программу «под себя» компании имеющей свой штат разработчиков будет намного дешевле, надёжнее и всегда можно будет что-то доработать, в конечном итоге, программа будет функционировать, соблюдая все нюансы и детали присущие конкретной компании, что даст несомненный плюс при работе.

## 1.3 Анализ функциональности представленных решений

Функционал перечисленных программ делится на 2 типа:

## 1.3.1 Полезный функционал для поиска сотрудников

* Автоматическое размещение вакансий на сайтах поиска работы.
* Автоматический сбор откликов с размещенных вакансий.
* Поиск резюме по заданным критериям
* Отображение статистики по сбору/откликам
* Возможность импортировать резюме из популярных программ, таких как Excel, Word и другие.

## 1.3.2 Вспомогательный функционал для решения кадровых проблем компании

* Ведение базы данных не только новых сотрудников, но и всех сотрудников компании.
* Автоматическое тестирование соискателей.
* Автоматический ответ соискателям по электронной почте.
* Автоматическая система расчета заработной платы.
* Оптимальная модель управления персоналом.
* Ведение персональных карточек сотрудников компании.

Функционал многих решений действительно имеет внушительные размеры. При правильной настройке эти функции позволяют значительно сократить время и затраты на поиск и управление сотрудниками.

Стоит заметить, на сегодняшний день программы стараются охватить всё больше сфер деятельности. Это значит, что уже довольно сложно найти программу, которая будет просто искать сотрудников. Зачастую программы включают в себя сторонний функционал, позволяющий решать второстепенные проблемы компании. Локальная необходимость этого функционала не учитывается производителем, а значит приходится переплачивать за функции, которые вероятно не будут использоваться.

В данной работе будет реализован функционал по сбору вакансий из открытых интернет источников, поскольку задачей является непосредственно анализ возможностей реализации автоматизированного поиска.

Поставку задачи в кейсе сайта профстажировки.рф выполнило предприятие МАГНИТ:

**Задача:** В процессе решения данного кейса необходимо разработать логику работы инструмента, способного осуществлять автоматизированный поиск резюме на различных работных сайтах по заданным параметрам. Описать функциональные и не функциональные требования к данному инструменту и исходя из этих требований определить стек технологий необходимых для его разработки и внедрения. При описании требований необходимо учесть возможность изменения процессов поиска и дальнейшего развития инструмента. Так же в рамках кейса необходимо провести анализ рынка ПО на предмет наличия подобных инструментов и проработать плюсы и минусы покупки или собственной разработки подобного решения.

**Результат:**

Вариант 1:

Подготовить презентацию на несколько слайдов, в которой будут описаны требования, логика работы и результаты анализа решений по автоматизированному поиску резюме.

Вариант 2:

Подготовить прототип ПО, позволяющий осуществлять подобный поиск.

## Стек технологий для разработки и внедрения автоматизированного поиска резюме на различных работных сайтах по заданным параметрам

## 2.1 Технологии Web Scrapping

Web Scrapping – это сбор данных с сайта и последующее преобразование их в необходимый формат. С развитием технологий появилась масса интернет ресурсов, которые несут в себе полезную информацию. Можно полагаться на какой-либо ресурс и пользоваться предоставляемыми им данными, а можно анализировать несколько ресурсов и получать больше качественной информации. Поиск резюме – это отличный пример для использования веб скраппинга, ведь вся работа сводится к следующему:

1. Ввод необходимых данных для поиска.
2. Автоматический поиск и сбор данных с сайтов.
3. Анализ данных и последующая их обработка (вывод либо действие).

Для того, чтобы разобраться со «Скраппингом», необходимо знать следующие понятия:

1. GET запрос – это обычный HTTP или HTTP/s запрос на URL ссылку с «Открытыми» параметрами, то есть все параметры (если они есть) запроса передаются в самой ссылке. Такие запросы используются для получения данных, которые может увидеть любой пользователь, и которые не обязательно скрывать.
2. POST запрос – это HTTP или HTTP/s запрос на URL ссылку с «Закрытыми» параметрами, то есть все параметры запроса передаются в теле запроса в скрытом виде. Такие запросы используются для авторизации на сайте, осуществления каких-либо действий на сайте, которые требуют дополнительных параметров. Например, «Поиск» требует наличия ключевых слов, так же на некоторых сайтах неавторизированным пользователям поиск запрещен, авторизация проверяется исходя из заголовка и тела нашего запроса. «Отправка комментария» требует нескольких параметров, зачастую это наличие авторизации, текст комментария, заголовок, время и так далее. Пожалуй, самый важный пример использования POST запроса – это авторизация. Существует огромное количество сайтов, на которых для того, чтобы получить доступ ко многим данным необходимо авторизоваться, и процесс веб скраппинга без авторизации становится невозможным.
3. «Парсинг» - это уже анализ и обработка полученных с помощью GET или POST данных. Зачастую это строка с HTML кодом, в котором необходимо получить только определенную часть. Используя инструменты языка программирования можно осуществить быстрый и качественный сбор информации.

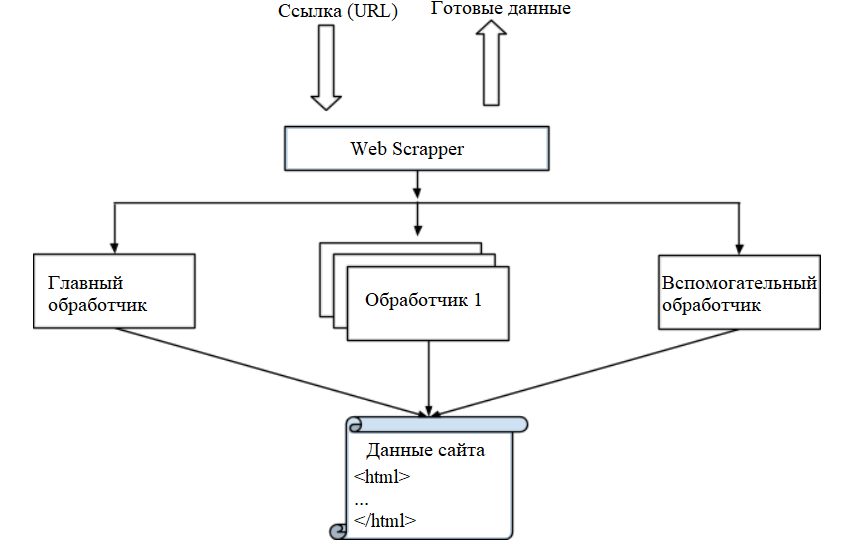
Сам Web Scrapping может быть реализован практически на любом языке программирования, разница будет лишь в реализации ключевых методов. Типовой вариант функциональной схемы работы скреппера приведен на рис.1.  


Рисунок.1 Функциональная схема работы скреппера

Через запросы получают данные с сайта с указанным URL. Далее проверяется есть ли среди обработчиков «Обработчик 1» … «Обработчик N» тот, который настроен под конкретный сайт. Если да, то используется он, иначе используется «Главный обработчик». Если по какой-либо причине «главному обработчику» не удаётся распознать данные, используется «Вспомогательный обработчик», который уже настраивается персонально, например администратором. На выходе получаем готовые данные, которые можем отобразить либо использовать их для дальнейших действий, таких как сохранение в Таблицу, или отправка ответа на почту и так далее.

## 2.2 Проблемы автоматизированного поиска

1. Ни один сайт не имеет идеальной верстки с точки зрения догматов веб-дизайна. Именно это делает каждый сайт уникальным и привлекательным для пользователей.
2. Почти каждый веб-разработчик пишет код под себя или просто, как умеет. Далеко не всегда код получается грамотным и качественным. Зачастую в нем можно найти огромное количество ошибок. В том числе грамматических. Все это делает «самописный» код абсолютно нечитаемым для скрапперов.
3. Масса веб-ресурсов использует HTML5, где каждый элемент может быть абсолютно уникальным.
4. Некоторые ресурсы содержат разнообразные защиты от копирования данных, а значит и от скрапинга. Это выражается в многоуровневой верстке, использовании JavaScript для рендеринга контента, проверки user-agent и т.д.
5. В зависимости от сезона или тематики целевого материала на сайте могут быть использованы разные макеты. Периодически это касается даже типичных страниц (сезонные акции, премиум статьи и т.д.).
6. Кроме полезных блоков, веб-страница часто изобилует “мусором” в виде рекламы, комментариев, дополнительных элементов навигации и т.д.
7. Исходный код может содержать ссылки на одни и те же картинки разных размеров, например — для превью.
8. Сайт может определить страну, в которой находится ваш сервер и отдать информацию на другом языке или вообще другую информацию.
9. У всех сайтов может быть разная кодировка, которая не отдается в ответе на запрос.

## 2.3 Обзор библиотеки BeautifulSoup.

Данная библиотека была разработана для упрощения веб скреппенга. Она позволяет сократить огромное количество времени и строк кода, для поиска необходимых данных со страницы. Библиотека работает на языке Python и позволяет анализировать и доставать данные с HTML и XML форматов страниц. Для установки библиотеки используется команда: pip install bs4.

Пример работы библиотеки:

from bs4 import BeautifulSoup

from urllib.request import urlopen

with urlopen('https://domkadrov.ru') as response:

soup = BeautifulSoup(response, 'html.parser')

for anchor in soup.find\_all('a'):

print(anchor.get('href', '/'))

Этот пример позволяет достать все ссылки с главной страницы сайта Дом Кадров.

## Инструменты диагностики программ для WEB

Для выявления проблем, а также для дальнейшего предотвращения ошибок, необходимо уметь получать и читать HTML код страницы. Зачастую на современных сайтах страницы подгружаются динамически, и порой сложно уловить малейшие изменения в реальном времени. Так же для осуществления авторизации, поиска и дополнительных возможностей сайта, необходимо понимать из чего состоит запрос. В исходном коде можно упустить некоторые параметры, либо они могут быть просто скрыты. Для этого необходимо смотреть не только полученный в результате запроса HTML код страницы, но и анализировать сам запрос. Нередко заголовок запроса состоит из множества параметров. И большинство из этих параметров являются обязательными для отправки, иначе сервер просто вернёт ошибку. Некоторые параметры современные сетевые библиотеки способны устанавливать автоматически, например параметр Content-Type сообщает клиенту какой будет тип передаваемого контента, и многие сетевые библиотеки умеют автоматически определять тип. Так же заголовки хранят ссылку, которая необходима для отправки запроса, а параметр Locale хранит ссылку, на которую нужно перейти после отправки запроса (если она имеется), так называемый редирект.

## 3.1 Анализ запроса с использованием функционала браузера Mozilla Firefox

В браузере Mozilla Firefox по умолчанию установлена так называемая «консоль разработчика» изображенная на рисунке 2. Доступ к консоли можно получить либо через нажатие правой кнопки мыши по области страницы, либо через сочетание клавиш Ctrl+Shift+E.

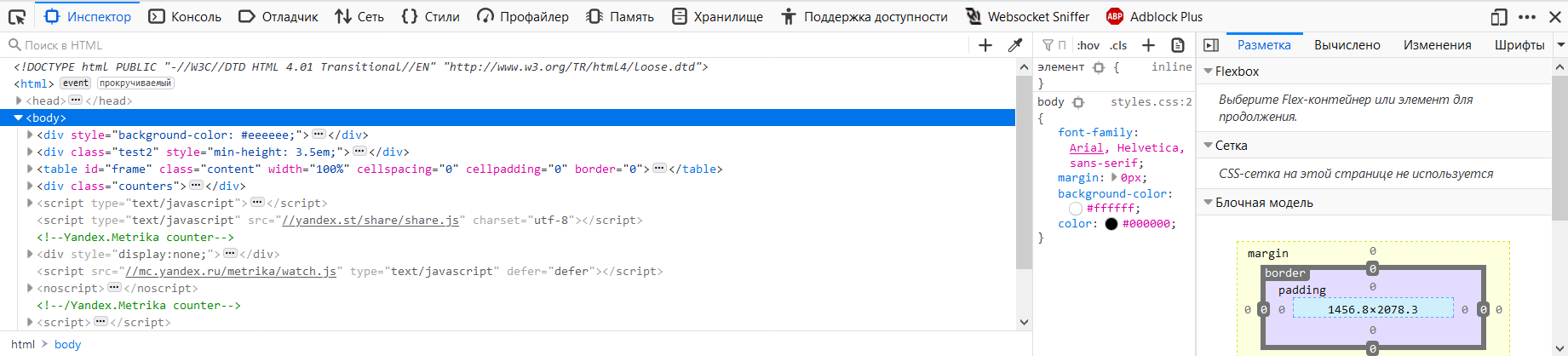


Рисунок 2. Консоль разработчика Mozilla Firefox

На вкладке Инспектор, изображенной на рисунке 2, мы можем наблюдать исходный код текущей страницы, а также при наведении на любой участок кода мы можем в реальном времени наблюдать к какому объекту страницы он относится. Так же этот код можно менять, результат мгновенно отобразится на странице, что в свою очередь полезно при веб верстке. Наибольший интерес представляет вкладка Сеть изображенной на рисунке 3.

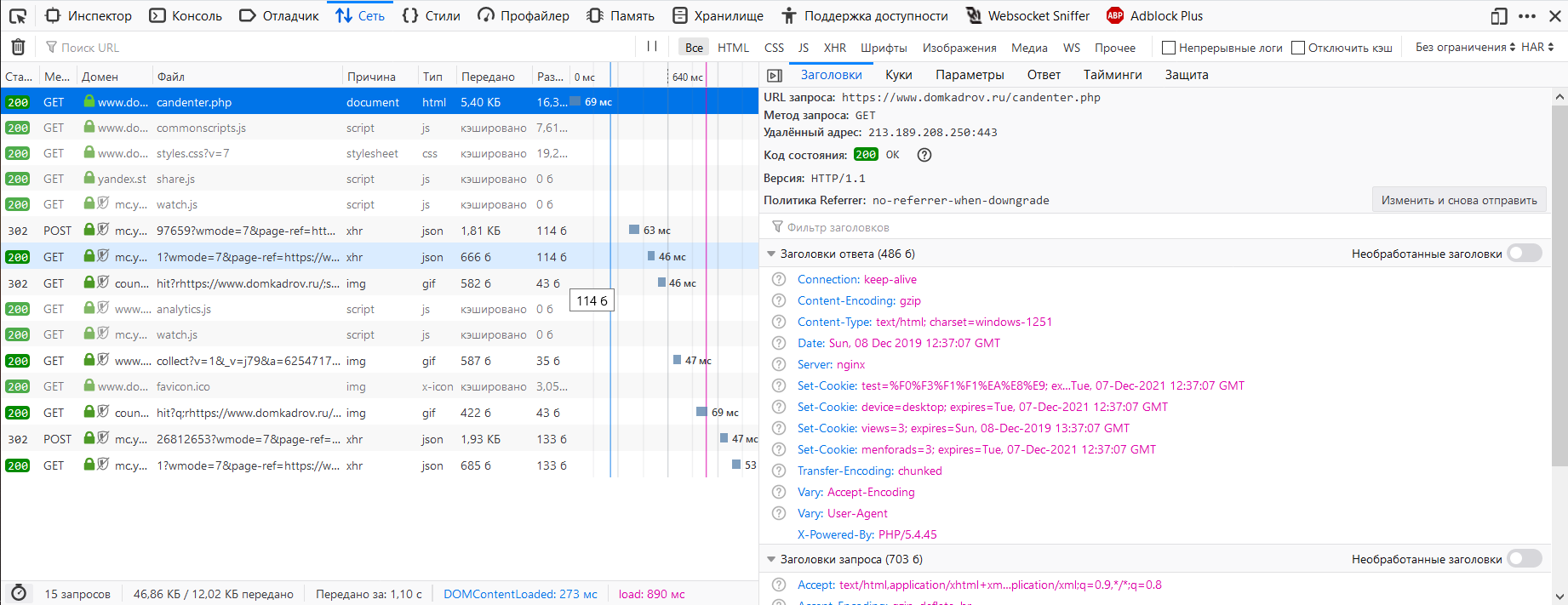


Рисунок 3. Вкладка «Сеть» на консоли разработчика Mozilla Firefox

В левой части консоли наблюдаются все запросы, которые были совершены при переходе на текущую страницу. Если нажать на любой запрос, то в правой части консоли откроется инструмент, позволяющий получить всю необходимую информацию о запросе. Информация о запросе поделена на вкладки:

* Заголовки – хранят информацию о заголовках запроса и заголовках ответа
* Куки – показывают всю сессионную информацию сайта
* Параметры – это те данные, которые передаются вместе с запросом
* Ответ – То, что сайт возвращает (часто это html код или JSON объект)

## 3.2 Анализ запросов с использованием программы HTTP Analyzer v7

Программа позволяет смотреть все запросы совершенные в системе. Любой системный процесс, отославший сетевой запрос, отлавливается и отображается пользователю, например на рисунке 4 видно, как процесс firefox.exe (браузер Mozilla Firefox) отослал несколько запросов. Эти запросы как раз были отосланы в момент открытия страницы сайта Дом Кадров.

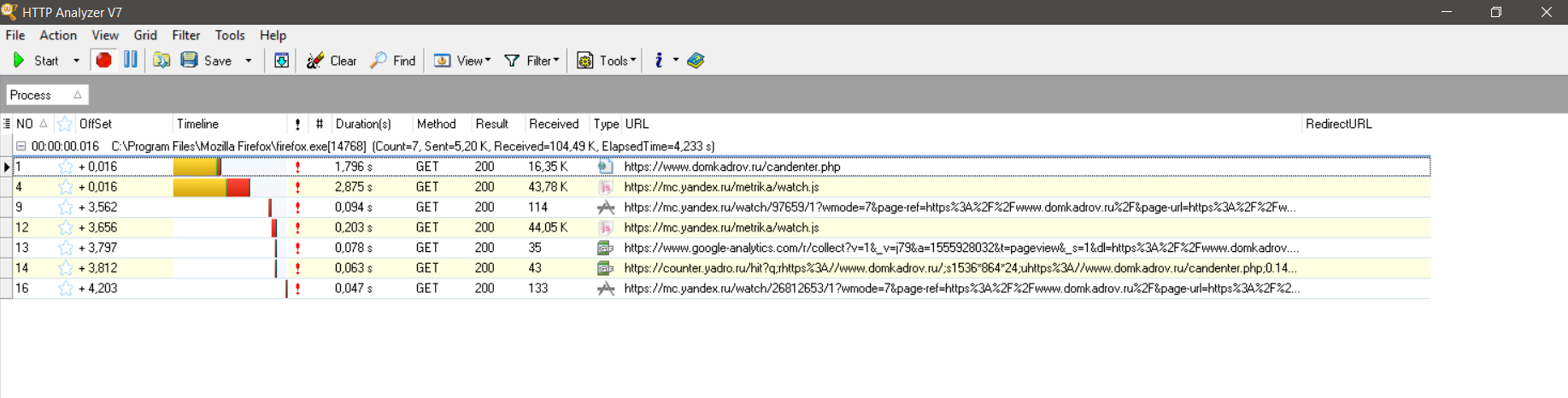


Рисунок 4. Отображение запросов в программе HTTP Analyzer v7.

Кликнув по любому запросу в нижней части программы открывается окно, отображенное на рисунке 5, позволяющее подробнее изучить запрос.

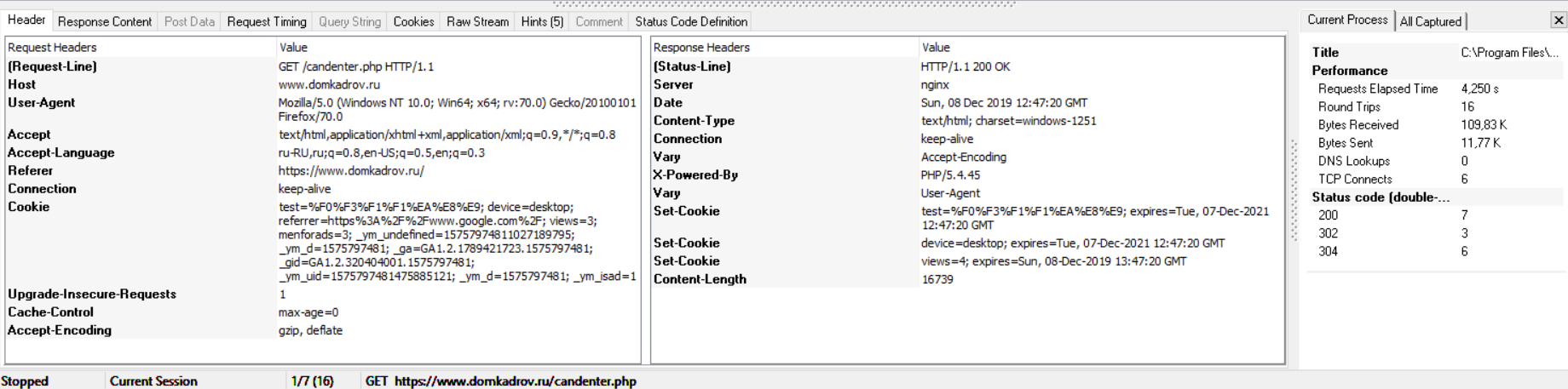


Рисунок 5. Окно с информацией о запросе в программе HTTP Analyzer v7.

Работа с этим окном аналогична с работой вкладки «Сеть» в консоли разработчика браузера Mozilla Firefox. Основное отличие этой программы от консоли разработчика браузера Mozilla Firefox в том, что программа HTTP Analyzer v7 позволяет просматривать запросы и ответы отправленные в любом процессе системы, в том числе и в разрабатываемой программе по поиску сотрудников. А значит намного упрощая процесс отладки и изучения запросов, ведь наблюдение за ними происходит в режиме реального времени и не происходит траты времени на вывод данных в самой программе.

## Программа для поиска резюме

Программа автоматизированного поиска резюме на работных сайтах была реализована на языке C++ в фреймворке Qt. Минусы этого решения: невозможность работы удалённо, использовались грубые методы добычи информации с сайтов, ибо библиотек веб скреппинга для C++ нет. Плюсы этого решения: скорость работы, простота отладки, простой интерфейс, после того как решение начинает стабильно работать, его не сложно перенести на Python и установить на сервер.

Внешний вид программы изображен на рисунке 6. Программа позволяет собирать резюме по фильтрам "профессия" и "город" с двух сайтов: Дом кадров и Авито.

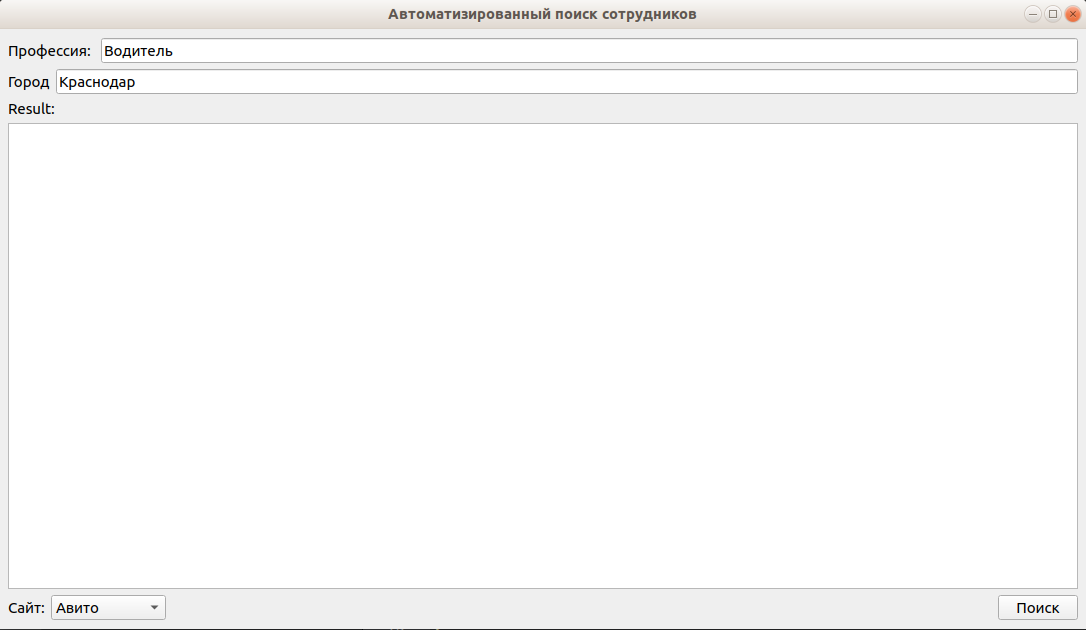


Рисунок 6. Внешний вид программы

В результате анализа популярных сайтов для вакансий были отобраны именно эти сайты, поскольку именно на них удалось бесплатно получить актуальные резюме. Остальные сайты, такие как hh.ru, rabota.ru, job.ru и так далее, требуют либо обязательного наличия зарегистрированной компании, либо приобретения доступа к базе, что значительно усложняло разработку и тестирование программы.

## 4.1 Поиск резюме на сайте domkadrov.ru

Во время разработки поиска резюме на сайте Дом Кадров были обнаружены следующие проблемы:

* Для поиска резюме обязательно необходимо быть авторизированным пользователем.
* Исходный код страницы оставляет желать лучшего и скреппинг превратился в настоящее испытание.

Проблему авторизации легко удалось решить. Для этого перед началом отправки запроса на поиск резюме, необходимо отправить запрос на авторизацию, который был обнаружен в консоли разработчика и изображен на рисунке 7.

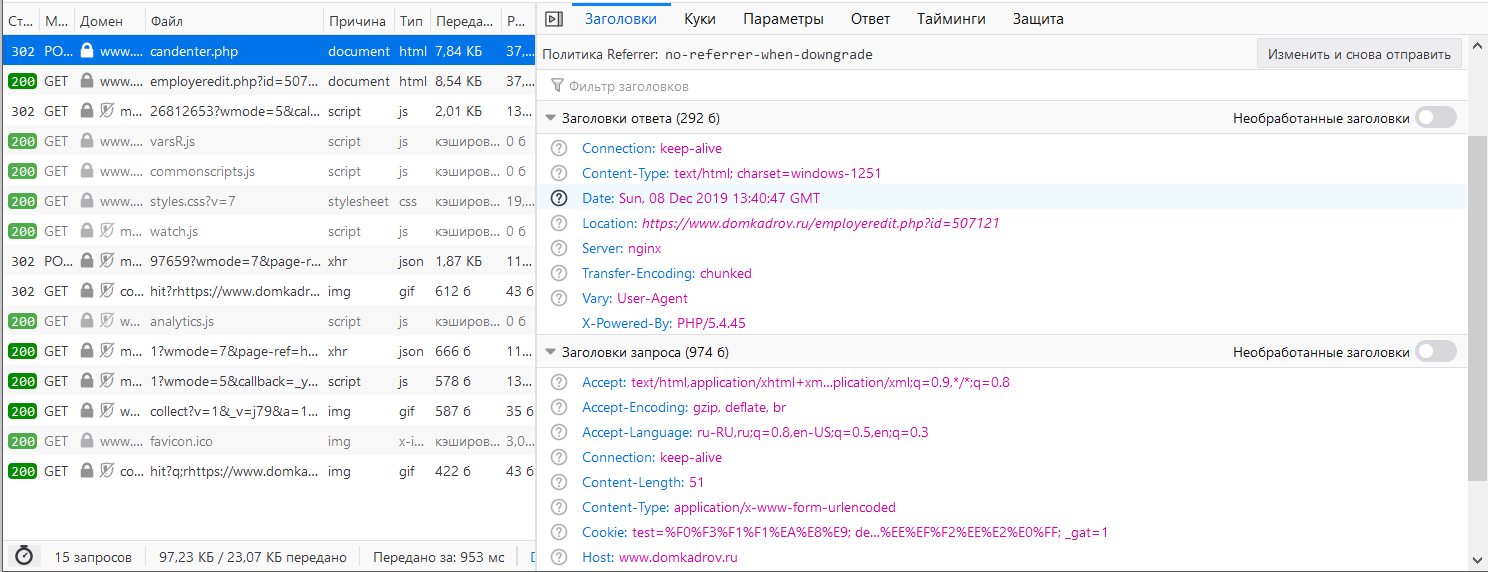


Рисунок 7. Запрос на авторизацию на сайт Дом кадров

После отправки точно такого же запроса как на рисунке в ответ приходят параметры Set-Cookie, устанавливающие в браузер куки об успешной авторизации на сайте, и Locale для редиректа на страницу нашего профиля. Наличие этих параметров равносильно успешной авторизации, а значит можно продолжить поиск. Загрузив страницу с необходимыми параметрами, через всё ту же консоль разработчика удаётся наблюдать ход отображения необходимых данных. И с помощью написанной функции QString Pars(QString str, QString x1,QString x2) , возвращающей строку в тексте str, находящуюся между строкой x1 и строкой x2, удалось собрать необходимые данные. Результат работы сбора резюме с сайта Дом Кадров показан на рисунке 8.

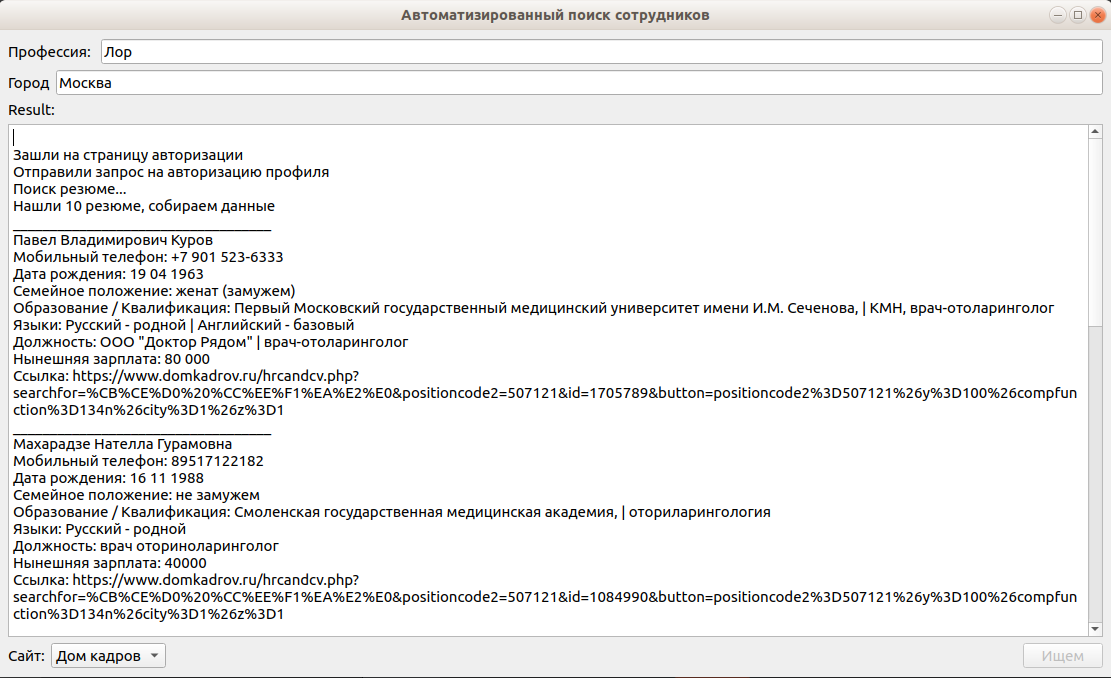


Рисунок 8. Результат работы по поиску резюме на сайте Дом Кадров.

## 4.2 Поиск резюме на сайте avito.ru

При работе с сайтом Авито масштабных проблем не возникло. Код понятный и читаемый, парсить его легко. Проблема заключается в невозможности связи с владельцами резюме по причине отсутствия авторизации. И даже после авторизации сайт просит оплатить подписку для возможности отображения номера. Поэтому было принято решение находить только основную информацию о кандидате и ссылку на его резюме. Достаточно просто загрузить главную страницу Авито, далее собрать необходимые параметры, изображенные на рисунке 9.

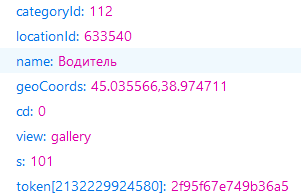


Рисунок 9. Параметры для успешного поискового запроса на сайте Авито.

После успешного сбора параметров можно отправлять запрос на поиск резюме. Если параметры действительные, то результатом будет список кандидатов. Разбиваем html код на Div-блоки. И отображаем эти блоки в виджете QTextEdit в формате html. Результат изображен на рисунке 10.

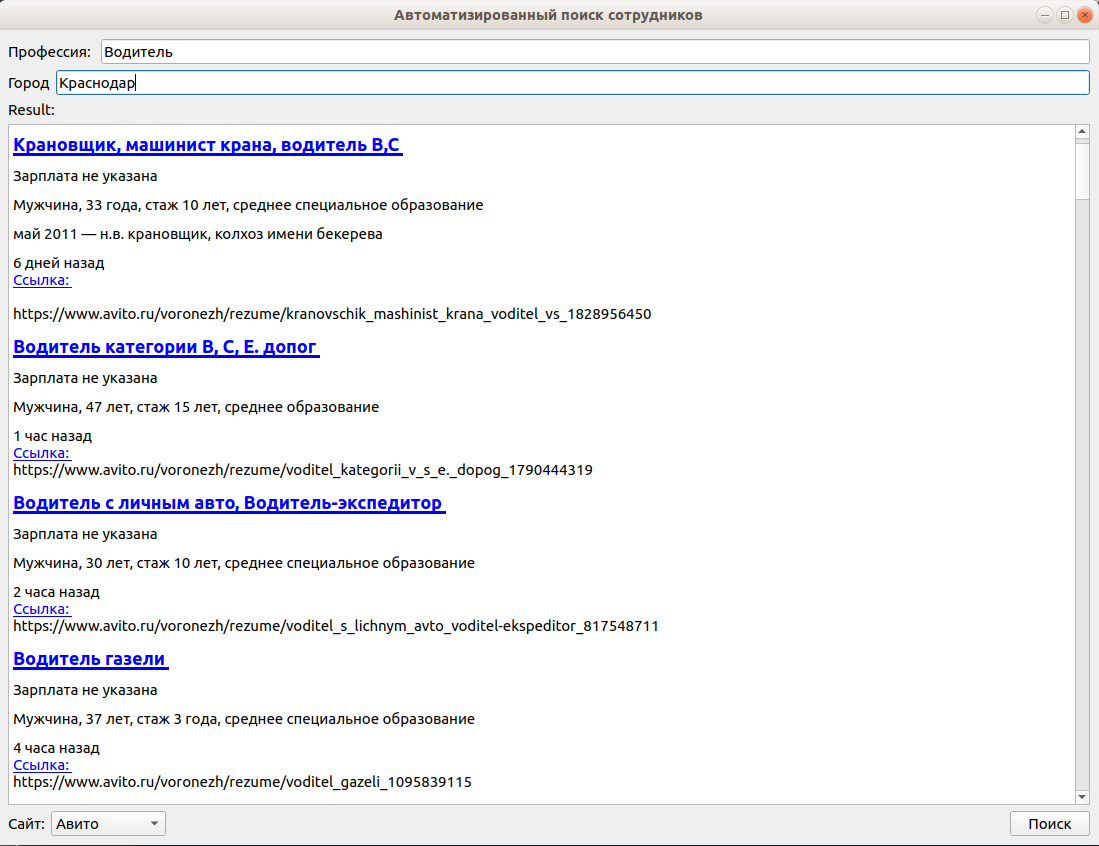


Рисунок 10. Результат работы по поиску резюме на сайте Авито.

Дальнейшее развитие проекта и подключение сторонних сайтов напрямую зависит от желания автора кейса с сайта профстажировки.ру.

## ЗАКЛЮЧЕНИЕ

В ходе разработки приложения для автоматизированного поиска резюме на различных работных сайтах по заданным параметрам были выполнены все задачи, поставленные в кейсе:

1. Разработана логика работы инструмента, способного осуществлять автоматизированный поиск резюме на сайтах для поиска работы.
2. Описаны функциональные и нефункциональные требования к данному инструменту.
3. Определён стек технологий, необходимых для реализации инструмента.
4. Проведен анализ рынка ПО на предмет наличия подобных инструментов.
5. Проработаны плюсы и минусы покупки или собственной разработки подобного решения.

Кроме того, были получены навыки в веб скреппинге. Подробно изучена структура запросов. Были изучены вспомогательные инструменты для анализа запросов и веб среппинга такие, как «Консоль разработчика» в браузере Mozilla Firefox, HTTP Analyzer v7, библиотека для веб скреппинга BeautifulSoup. Была изучена структура современных сайтов, а также был разработан и апробирован прототип программы по автоматизированному сбору информации из открытых источников Интернет.

## ЛИТЕРАТУРНЫЕ ИСТОЧНИКИ

1. Кондукова Е. Qt 5.3 Профессиональное программирование на C++// Издательство: BHV, 2017 – 928с

2. А.А. Дубаков СЕТЕВОЕ ПРОГРАММИРОВАНИЕ// Издательство: НИУ ИТМО, 2013 - 248 с

3. Документация на русском языке. [Электронный ресурс] Документация, URL:http://wiki.python.su/%D0%94%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%B0%D1%86%D0%B8%D0%B8/BeautifulSoup [Дата обращения: 18 октября 2019]

4. 9 Программ для поиска сотрудников. [Электронный ресурс] Статья, URL: https://kontur.ru/articles/1579 [Дата обращения: 29 ноября 2019]

5. Web Scrapping с помощью python. [Электронный ресурс] Статья, URL: https://habr.com/ru/post/280238/ [Дата обращения: 25 октября 2019]

6. Web Scrapping с помощью python. [Электронный ресурс] Статья, URL: https://4brain.ru/blog [Дата обращения: 28 октября 2019]