

DOI: 10.32517/0234-0453-2022-37-5-26-37

# ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ О ВАКАНСИЯХ ДЛЯ ВЫЯВЛЕНИЯ АКТУАЛЬНЫХ ПОТРЕБНОСТЕЙ РЫНКА ТРУДА

Ю. А. Морозова<sup>1</sup> ✉<sup>1</sup> *Национальный исследовательский университет «Высшая школа экономики», г. Москва, Россия*✉ [yumorozova@hse.ru](mailto:yumorozova@hse.ru)

## Аннотация

В результате процессов цифровизации производства и перехода на новые технологии на рынке труда изменяются требования к навыкам и умениям специалистов. Наличие рабочей силы требуемой квалификации критически важно для стимулирования экономического роста. В октябре 2021 года Правительством Российской Федерации был утвержден федеральный проект «Развитие кадрового потенциала ИТ-отрасли» в составе национальной программы «Цифровая экономика Российской Федерации», который реализуется в рамках перечня инициатив социально-экономического развития Российской Федерации до 2030 года. В условиях изменяющихся технологических трендов одной из основных становится задача своевременной адаптации образовательных программ высших учебных заведений к подготовке специалистов, соответствующих потребностям рынка труда. В статье предлагается инструментальный интеллектуальный анализ данных о вакансиях для выявления актуальных навыков и умений, требуемых на рынке труда. На примере рынка ИТ-труда на основе данных, собранных из сервиса hh.ru, были проанализированы ключевые навыки, указанные в размещенных на сайте вакансиях, а также выполнена кластеризация вакансий. Предложенный подход может использоваться на регулярной основе образовательными учреждениями для актуализации учебных планов образовательных программ.

**Ключевые слова:** интеллектуальный анализ данных, рынок труда, вакансии, навыки ИТ-специалистов, образование.

## Для цитирования:

Морозова Ю. А. Интеллектуальный анализ данных о вакансиях для выявления актуальных потребностей рынка труда. *Информатика и образование*. 2022;37(5):26–37. DOI: 10.32517/0234-0453-2022-37-5-26-37

---

# DATA MINING OF VACANCY DATA TO IDENTIFY THE CURRENT LABOR MARKET NEEDS

Yu. A. Morozova<sup>1</sup> ✉<sup>1</sup> *National Research University Higher School of Economics, Moscow, Russia*✉ [yumorozova@hse.ru](mailto:yumorozova@hse.ru)

## Abstract

As a result of manufacturing digitization and the transition to new technologies, the labor market requirements for specialists' skills and abilities are changing. Employing an appropriately qualified workforce is critical to stimulating economic growth. In October 2021, the federal project "Development of the IT industry human resources" of the national program "Digital Economy of the Russian Federation" was approved by the Government of Russia. The project is being implemented as part of the list of initiatives for the socio-economic development of the Russian Federation until 2030. In the context of changing technological trends, it is crucial to adapt the educational programs of higher educational institutions in timely manner to ensure training of such specialists who meet the demands of the labor market. The article proposes a data mining toolkit to analyze vacancy data in order to identify relevant skills and abilities required in the labor market. Based on the IT labor market data collected from the hh.ru service, I analyzed the key skills indicated in the vacancies posted on the site. Additionally, vacancy clustering was performed. Educational institutions can regularly use the suggested approach to update educational program curricula.

**Keywords:** data mining, labor market, vacancies, IT skills, education.

## For citation:

Morozova Yu. A. Data mining of vacancy data to identify the current labor market needs. *Informatics and Education*. 2022;37(5):26–37. (In Russian.) DOI: 10.32517/0234-0453-2022-37-5-26-37

## 1. Введение

Область информационных технологий в настоящее время является самой быстрорастущей. Каждый год в ней возникают новые направления, технологии и, соответственно, рабочие места. Но, несмотря на доступность множества различных онлайн-курсов, публикаций и образовательных программ, позволяющих повысить свою квалификацию, рынок ИТ-труда страдает от нехватки специалистов.

Новые должности требуют знания не только новых технологий, но и основ профессии. Проблема в том, что краткосрочные курсы направлены на предоставление дополнительного образования, в них практически не уделяется внимания основам и смежным дисциплинам. В то же время колледжи и университеты сталкиваются с другой проблемой: они не успевают вовремя меняться, чтобы соответствовать потребностям рынка [1]. Кризис, вызванный пандемией и сложной политической ситуацией, нанес ущерб российской экономике, но также он стимулировал развитие рынка ИТ-услуг [2]. Возникают новые деловые инициативы, реализуются организационные изменения и открываются новые возможности для роста. Все эти факторы обуславливают актуальность проблемы мониторинга ситуации на рынке труда и потребность в инструментах, позволяющих на основе анализа больших данных отслеживать изменения потребностей в специалистах.

Попытки исследовать изменения на рынке труда, выявить актуальные требования к квалификации работников, классифицировать появляющиеся профессии время от времени предпринимались исследователями в разных странах, но они носят нерегулярный характер [3–8].

В последние десятилетия интернет (веб-ресурсы) стал важной частью процессов найма и поиска работы [9]. Онлайн-порталы, которые собирают вакансии в интернете, стали для бизнеса и лиц, ищущих работу, важным источником информации о спросе и предложении, а также инструментом для подбора персонала инновационными способами. Кроме того, эти ресурсы открывают новые области исследований со значительным потенциалом для решения широкого круга задач, особенно тех, которые связаны с использованием новых источников данных и аналитических методов для изучения рынка труда. Огромное количество вакансий, доступных в настоящее время на онлайн-порталах, позволяет получить точную и подробную информацию о динамике и тенденциях рынка труда в интернете [10]. Это указывает на то, что веб-источники могут быть хорошей базой данных для анализа вакансий.

Ряд исследований был посвящен анализу объявлений о вакансиях. Более ранние работы главным образом основывались на печатных объявлениях о вакансиях [11–13], в недавних анализируются онлайн-вакансии. В работах [14, 15] авторы исследуют компетенции, требуемые от работников в Словакии, Чешской Республике, Дании и Ирландии для низ-

коквалифицированных и среднеквалифицированных профессий в различных секторах. Сравнивается спрос на формальную квалификацию, когнитивные и некогнитивные навыки. В работе [16] используется выборка вакансий для профессий средней квалификации, полученная из журнала *Burning Glass*, для изучения явления повышения требований к квалификации во время Великой рецессии в США. Авторы [17] на основе выборки из 2 млн объявлений о вакансиях, опубликованных в интернете, оценивают, какие образовательные, профессиональные и другие требования наиболее часто предъявляются американскими работодателями.

По сравнению с традиционными источниками веб-данные подробнее и предоставляют информацию, которая, возможно, ранее была недоступна. Онлайн-источники также позволяют исследователям собирать более крупные и разнообразные наборы данных. Другим преимуществом является гибкий, быстрый, простой и дешевый процесс сбора данных. Тем не менее существуют некоторые ограничения в отношении онлайн-данных о вакансиях, которые исследователи должны учитывать [15, 18]. Во-первых, не обо всех вакансиях информация публикуется открыто и не все вакансии соответствуют реальным должностям. Некоторые ведущие компании предпочитают подбирать кандидатов на ключевые должности посредством частных связей и рекомендаций. Более того, исследователи [18] пришли к выводу, что онлайн-объявления о вакансиях в основном ориентированы на высококвалифицированных кандидатов, ищущих должности «белых воротничков», а также на специалистов высокого уровня в области науки, технологии, инженерии и математики. Однако для нашего исследования, результаты которого ориентированы на высшие учебные заведения, это не является ограничением.

При анализе вакансий применяются как статистические методы, так и методы машинного обучения, обработки естественного языка (NLP – Natural Language Processing) [19–21]. В частности, методы интеллектуального анализа текстов были использованы для определения ландшафта знаний, необходимых организациям индустрии 4.0, в исследовании [22]. Авторам удалось выявить две группы профилей вакансий: ориентированные исключительно на индустрию 4.0 и ориентированные на менеджмент и информатику.

Алгоритм кластеризации вакансий на основе требований к навыкам применен в работе [23]. В результате обработки порядка 3 млн объявлений с сервиса *hh.ru* удалось выявить 15 кластеров вакансий, включая маркетинг, техническое обеспечение, большие данные, программное обеспечение, администрирование, безопасность, веб-дизайн, инженерия, аналитика, «мягкие» навыки (*soft skills*), менеджмент, тестирование, ERP (*enterprise resource planning*).

Целью данной статьи является выработка общего подхода к анализу данных о вакансиях для мониторинга и отслеживания изменений на рынке труда,

выявления актуальных профессий и квалификационных требований в целях обеспечения своевременной адаптации учебных планов образовательных программ высших учебных заведений.

## 2. Методы

### 2.1. Набор данных

В качестве источника данных для исследования был выбран популярный сервис для поиска работы в России HeadHunter (hh.ru). HeadHunter предоставляет API (Application Programming Interface) для получения вакансий и подробную документацию по API, что позволяет легко собирать необходимые данные в автоматическом режиме.

Для сбора данных были выбраны следующие параметры: идентификатор вакансии, имя работодателя, название вакансии, оклад или диапазон оклада, требуемый опыт, тип графика работы, теги навыков, описание вакансии, область вакансии, профессиональные роли вакансии, специализация вакансии, профессиональная область вакансии, время публикации вакансии.

Сбор и анализ данных осуществлялись с использованием языка программирования Python. Для работы с запросами API использовался пакет Python «Requests». Для анализа данных использовался пакет Python «Pandas». Пакет Python «Re» использовался для очистки данных из HTML-тегов.

В результате процедуры сбора данных с сайта hh.ru были загружены все вакансии в области информационных технологий в Москве и Санкт-Петербурге за период с 25.10.2021 года по 02.12.2021 года. В выборке содержится информация о 47 330 уникальных вакансиях. Все зарплаты в валютах, отличных от рубля, были пересчитаны в рубли по курсу Центрального банка России на 02.12.2021 г.

### 2.2. Методы анализа данных о вакансиях

Первым шагом исследования вакансий является разведочный анализ, цель которого — быстро определить основные характеристики вакансий, востребованных на рынке труда:

- список работодателей, вакансии которых наиболее часто встречаются в выборке;
- наименования вакансий, которые наиболее часто встречаются в выборке;
- распределение вакансий по заявленным графикам работы;
- распределение вакансий по требуемому опыту работы;
- средний разброс предлагаемого размера заработной платы для каждой должности;
- список наиболее востребованных работодателями навыков;
- наиболее востребованные навыки для каждой должности.

Наименования одной и той же должности в объявлениях о вакансиях могут иметь разные написа-

ния, поэтому результаты разведочного анализа имеют ограничения. Поскольку количество уникальных наименований довольно велико, при отборе самых популярных из них из результатов анализа исключаются редкие наименования должностей, а это большая часть выборки. Возникает задача кластеризации вакансий, т. е. определения типового набора должностей с характерным набором требуемых навыков.

В данной статье предлагается кластеризация вакансий на основе матрицы частоты терминов (term frequency matrix), составленной из слов наименования должности и требуемых навыков. Предобработка и подготовка набора данных для кластеризации включала следующие операции:

- удаление служебных знаков и знаков пунктуации (пакет Python «Re»);
- токенизация по отдельным словам (пакет Python «Razdel»);
- лемматизация (пакет Python «Pymorphy2»);
- удаление стоп-слов на английском и русских языках (пакет Python «Stop\_words»);
- перевод всех терминов на английский язык, поскольку в объявлениях присутствуют наименования и на русском, и на английском языке (пакет Python «Googletrans»);
- преобразование вектора строк терминов в матрицу частот терминов (пакет Python «Scikit-learn»);
- фильтрация терминов, которые встречаются реже, чем в десяти вакансиях;
- фильтрация вакансий, в характеристиках которых после всех преобразований оказалось менее трех терминов.

Для определения количества кластеров использовались «метод локтя» (elbow method) и средний балл силуэта (silhouette method).

Для кластеризации были опробованы методы KMeans, Agglomerative Clustering, OPTICS, BisectingKMeans библиотеки Python «Scikit-learn». По критериям качества и скорости кластеризации был выбран KMeans.

## 3. Результаты

### 3.1. Разведочный анализ данных о вакансиях

В анализируемой выборке количество уникальных работодателей, предлагающих вакансии в сфере информационных технологий, составило 10 408, но среди них выделились лидеры по количеству вакансий — «Сбербанк», Ozon, «Яндекс», МТС, VK (табл. 1), из чего можно сделать вывод, что именно эти компании диктуют требования к специалистам на рынке и задают тренд. Доля вакансий, приходящаяся на 30 компаний-лидеров по количеству вакансий, составляет 16,7 % от общего числа вакансий в выборке.

Распределение вакансий по опыту работы (табл. 2) показывает, что около 60 % вакансий ориентировано на молодых специалистов, из них 10,8 % вакансий не требует опыта. Это можно объяснить быстро меняющимися требованиями и появлением

Таблица 1 / Table 1

**Список работодателей, вакансии которых наиболее часто встречаются в выборке**  
**List of employers whose vacancies are the most frequent in the sample**

Работодатель	Количество объявлений о вакансиях
«Сбербанк»	1491
Ozon	718
«Яндекс»	469
МТС	446
VK	368
«Газпром нефть»	347
«Билайн»	299
JCat.ru	287
«Иннотех», группа компаний	282
Банк ВТБ (ПАО)	263
«Тинькофф»	247
«Лига Цифровой Экономики»	204
ЛАНИТ	187
«Лаборатория Касперского»	172
ANCOR	169
JetBrains	166
«Газпром-медиа. Развлекательное телевидение» (ГПМ РТВ)	159
«Альфа-Банк»	150
«Ростелеком»	142
Skyeng	137
Bell Integrator	137
Selecty	133
«ЭР-Телеком»	132
«Газпромбанк»	131
ИК СИБИНТЕК	124
РОСБАНК, Societe Generale Group (Russia)	118
«Первый Бит»	115
LLC FINTERNA	114
КРОК	112
DINS	108

Таблица 2 / Table 2

**Распределение вакансий по опыту работы**  
**Distribution of vacancies by work experience**

Опыт работы	Процент объявлений о вакансиях
От одного года до трех лет	49,5
От трех до шести лет	36,4
Нет опыта	10,8
Более шести лет	3,3

новых технологий. Только в 3,3 % вакансий требуется опыт более шести лет.

По распределению вакансий по типам графиков работы (табл. 3) видно, что по-прежнему большая часть вакансий ориентирована на полный рабочий день (69,7 %), однако высокий процент вакансий предполагает и удаленную работу (22,9 %).

Список наиболее популярных вакансий (табл. 4) условный, так как в нем не учитывается разное написание одной и той же позиции. Но, как будет видно из последующего кластерного анализа, список дает адекватное представление о спектре востребованных на рынке позиций. Наиболее популярные позиции — системный администратор, программист 1С, системный аналитик, аналитик, бизнес-аналитик.

В пятерке наиболее часто встречающихся требований (табл. 5) — знание SQL, английского языка, Git, Python, Linux. При этом можно отметить, что среди требований можно встретить и знание определенных технологий (hard skills), и «мягкие» навыки (soft skills), такие как управление проектами, работа в команде, грамотная речь, аналитическое мышление.

Средний диапазон заработной платы для наиболее популярных наименований вакансий (табл. 6) рассчитан на основе данных тех вакансий, в которых он указан. Высокая заработная плата, предлагаемая в вакансии, может говорить о высоких требованиях к специалисту, а также о нехватке специалистов требуемой квалификации на рынке.

Таблица 3 / Table 3

**Распределение вакансий по типам графиков работы**  
**Distribution of vacancies by types of work schedules**

График работы	Процент объявлений о вакансиях
Полный день	69,7
Удаленная работа	22,9
Гибкий график	5,4
Сменный график	1,8
Вахтовый метод	0,2

Таблица 4 / Table 4

**Список наименований вакансий,  
наиболее часто встречающихся в выборке**  
**List of job titles most frequently found in the sample**

Наименование должности	Количество объявлений о вакансиях
Системный администратор	593
Программист 1С	533
Системный аналитик	515
Аналитик	267
Бизнес-аналитик	240
Специалист технической поддержки	237
Frontend-разработчик	196
Модератор (удаленно)	194
Java-разработчик	179
Технический писатель	149
DevOps инженер	144
Инженер технической поддержки	142
Менеджер по продажам	137
Интернет-маркетолог	122
Руководитель проектов	117
Менеджер проектов	103
DevOps Engineer	100
Сервисный инженер	99
IOS-разработчик	99
QA Engineer	97
PHP-разработчик	93
Монтажник слаботочных систем	81
Системный администратор Linux	78
Графический дизайнер	76
Руководитель проекта	76
Продавец	76
Аналитик данных	75
Инженер-программист	72
Data Engineer	71
Android разработчик	71

Таблица 5 / Table 5

**Список навыков,  
наиболее часто встречающихся в выборке**  
**List of skills most frequently found in the sample**

Требуемые навыки	Процент объявлений о вакансиях
SQL	15,3
английский язык	11,5
Git	10,5
Python	8,9
Linux	8,8
управление проектами	7,0
работа в команде	6,7
JavaScript	6,4
Java	6,4
PostgreSQL	5,8
грамотная речь	4,7
Atlassian Jira	4,6
MS SQL	3,6
Docker	3,5
ООП	3,3
HTML	3,3
MS PowerPoint	3,2
MySQL	3,2
аналитическое мышление	3,2
PHP	3,1
C#	3,1
ведение переговоров	2,8
деловая коммуникация	2,8
CSS	2,7
Adobe Photoshop	2,6
бизнес-анализ	2,5
пользователь ПК	2,3
1С программирование	2,3
C++	2,2
деловая переписка	2,2

Окончание табл. 5 / End of the table 5

Требуемые навыки	Процент объявлений о вакансиях
разработка технических заданий	2,2
работа с большим объемом информации	2,2
BPMN	2,2
Spring Framework	2,2
грамотность	2,1
деловое общение	2,1
REST	2,1
1С:предприятие 8	2,1
тестирование	2,1
React	2,1
Scrum	2,0
организаторские навыки	2,0
UML	2,0
Agile project management	2,0
TypeScript	2,0
XML	1,8
Kubernetes	1,8
Atlassian Confluence	1,7
креативность	1,7
информационные технологии	1,7

Таблица 6 / Table 6

**Средний диапазон заработной платы для наиболее популярных наименований вакансий**  
**Average salary range for the most popular job titles**

Наименование вакансии	Средний минимум заработной платы, руб.	Средний максимум заработной платы, руб.
Java-разработчик	150 000	239 000
Frontend-разработчик	129 000	193 000
Программист 1С	127 000	165 000
Системный аналитик	126 000	194 000
Бизнес-аналитик	104 000	136 000
Аналитик	97 000	128 000
Технический писатель	78 000	127 000
Системный администратор	69 000	83 000
Специалист технической поддержки	49 000	60 000
Модератор (удаленно)	26 000	32 000

Анализ требований к кандидатам позволил выявить наиболее востребованные навыки для наиболее популярных вакансий (табл. 7 и рис. 1).

В разведочном анализе фокус был сделан на наиболее популярных наименованиях вакансий. Более полное представление о спектре востребованных профессий и навыках дает кластерный анализ.

Таблица 7 / Table 7

**Наиболее востребованные навыки для популярных вакансий**

**Most sought-after skills for popular job titles**

Системный администратор		Системный аналитик		Аналитик	
Требуемые навыки	Процент вакансий	Требуемые навыки	Процент вакансий	Требуемые навыки	Процент вакансий
Active Directory	21,2	SQL	40,8	SQL	34,5
администрирование сетевого оборудования	20,7	UML	36,9	аналитическое мышление	22,1
администрирование серверов Windows	20,4	BPMN	31,5	MS PowerPoint	21,3
Linux	19,7	системный анализ	31,1	анализ данных	16,1
настройка ПК	16,0	Atlassian Jira	20,6	работа с большим объемом информации	16,1
настройка сетевых подключений	14,7	REST	19,6	BPMN	13,9

Окончание табл. 7 / End of the table 7

Системный администратор		Системный аналитик		Аналитик	
Требуемые навыки	Процент вакансий	Требуемые навыки	Процент вакансий	Требуемые навыки	Процент вакансий
Windows 7	14,0	разработка технических заданий	19,6	бизнес-анализ	12,7
TCP/IP	13,2	XML	18,6	английский язык	10,5
DHCP	12,6	бизнес-анализ	16,5	UML	10,5
настройка программного обеспечения	12,1	Atlassian Confluence	12,8	Python	10,1
администрирование	9,6	SOAP	12,8	разработка технических заданий	9,4
VMware	8,8	MS Visio	9,9	Atlassian Jira	9,0
настройка DNS	8,3	постановка задач разработчикам	8,3	Ms Visio	9,0
техническая поддержка	7,8	аналитическое мышление	8,2	MS SQL	9,0
администрирование серверов Linux	7,3	английский язык	7,6	работа в команде	7,9
Zabbix	6,6	API	7,4	Power BI	7,5
Windows Server 2003	6,6	управление проектами	6,8	управление проектами	7,5
сборка ПК	6,2	JSON API	6,2	аналитика	7,1
PowerShell	6,1	MS SQL	5,8	MS Excel	7,1
Hyper-V	6,1	моделирование бизнес-процессов	4,7	работа с базами данных	6,7

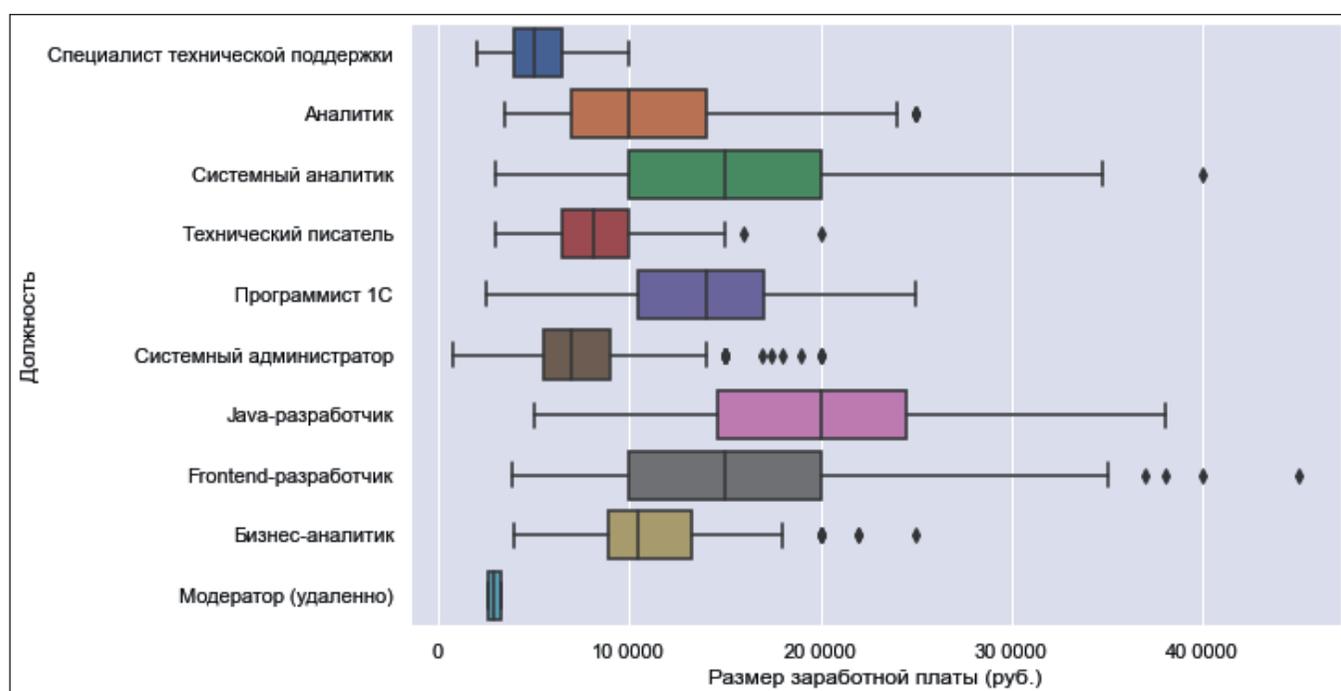


Рис. 1. Распределение заработной платы для наиболее популярных вакансий

Fig. 1. Salary distribution for the most popular job titles

### 3.2. Кластерный анализ данных о вакансиях

Для определения оптимального количества кластеров были рассчитаны метрики суммы квадратов расстояний наблюдений до их ближайшего центра кластера и средний балл силуэта (рис. 2). Однако графики этих метрик не позволили однозначно определить количество кластеров, при котором одновременно были бы представлены все направления специализаций, и кластеры существенно различались по содержанию, поэтому их количество было определено эмпирически.

В таблице 8 представлены 15 выявленных кластеров вакансий с наиболее типичными наименованиями позиций и требуемыми навыками.

В таблице 9 представлены средний разброс зарплат для каждого кластера вакансий, наиболее часто требуемый опыт работы и предлагаемый график работы. Наиболее высокий уровень заработных плат предлагается в вакансиях кластеров Java-разработки и разработки мобильных приложений. Для всех кластеров, кроме Java-разработки и управления проектами, наиболее часто требуемый опыт работы составляет от одного года до трех лет. Для Java-разработки и управления проектами наиболее востребован специалист с опытом работы от трех до шести лет. Для всех кластеров вакансий наиболее частый график работы — полный день, однако для специалистов, занимающихся разработкой, процент вакансий с полным рабочим днем наименьший, что говорит о наибольшем распространении удаленной работы и гибких графиков для разработчиков.

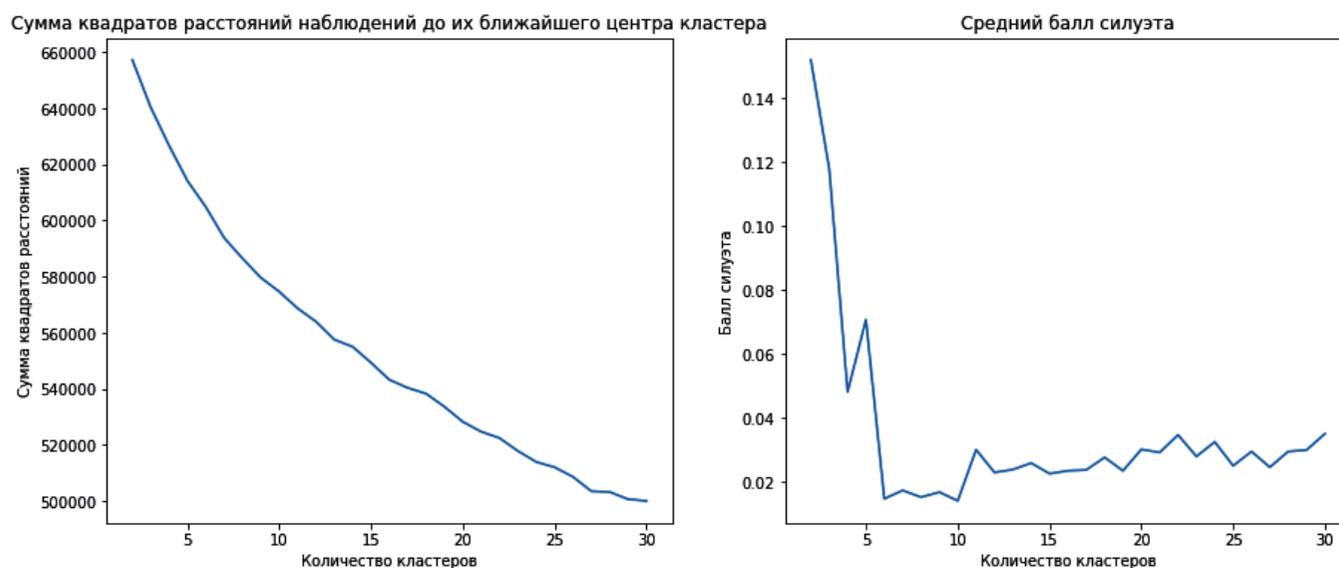


Рис. 2. Определение количества кластеров

Fig. 2. Determining the number of clusters

Таблица 8 / Table 8

#### Кластеры вакансий и требуемые навыки Vacancy clusters and required skills

Кластер	Название кластера	Наименования вакансий	Требуемые навыки
1	DevOps-инженерия	DevOps инженер, DevOps Engineer, системный администратор Linux, DevOps engineer, системный администратор	Linux, Python, Docker, Git, Bash, PostgreSQL, Kubernetes, Ansible, SQL, CI/CD, nginx, Jenkins, DevOps, Zabbix, MySQL, Java, администрирование серверов Linux, GitLab, Grafana, TCP/IP
2	Системное администрирование	Системный администратор, специалист технической поддержки, инженер технической поддержки, ведущий системный администратор, старший системный администратор	настройка ПК, настройка сетевых подключений, Linux, настройка ПО, Active Directory, администрирование сетевого оборудования, Windows 7, TCP/IP, техническая поддержка, администрирование серверов Windows, ремонт ПК, администрирование, информационные технологии, DHCP, сборка ПК, VMware, Help desk, администрирование серверов Linux, техническое обслуживание, Cisco
3	Администрирование и техническая поддержка	Технический писатель, системный администратор, бизнес-аналитик, специалист технической поддержки, аналитик	английский язык, SQL, Python, Linux, MS PowerPoint, Git, грамотная речь, Java, грамотность, работа в команде, C#, Adobe Photoshop, JavaScript, креативность, C++, Atlassian Jira, управление проектами, аналитическое мышление, C/C++, analytical skills

Продолжение табл. 8 / Continuation of the table 8

Кластер	Название кластера	Наименования вакансий	Требуемые навыки
4	Java-разработка	Java-разработчик, Java Developer, Java developer, Java разработчик, Senior Java Developer	Java, Spring Framework, Git, PostgreSQL, SQL, Hibernate ORM, Apache Maven, Apache Kafka, Spring Boot, rest, OOP, Docker, Kotlin, Linux, Java SE, Java EE, Spring, Junit, английский язык, MongoDB
5	Системный анализ и проектирование	Системный аналитик, аналитик, ведущий системный аналитик, бизнес-аналитик, аналитик бизнес-процессов	SQL, UML, BPMN, системный анализ, разработка технических заданий, бизнес-анализ, Atlassian Jira, XML, REST, Atlassian Confluence, MS Visio, постановка задач разработчикам, SOAP, моделирование бизнес-процессов, аналитическое мышление, управление проектами, MS SQL, английский язык, системное мышление, JSON API
6	Администрирование проектов	Менеджер проектов, специалист технической поддержки, администратор проектов, руководитель проектов, Project Manager	работа в команде, деловая коммуникация, деловое общение, грамотная речь, деловая переписка, английский язык, управление проектами, организаторские навыки, ведение переговоров, работа с большим объемом информации, пользователь ПК, MS PowerPoint, грамотность, ориентация на результат, телефонные переговоры, проведение презентаций, аналитическое мышление, MS Outlook, обучение и развитие, многозадачность
7	Техническая поддержка и сервис	Монтажник слаботочных систем, специалист технической поддержки, аналитик, технический писатель, системный администратор	грамотная речь, работа с большим объемом информации, пользователь ПК, грамотность, MS PowerPoint, удаленная работа, аналитическое мышление, английский язык, умение работать в команде, поиск информации в интернете, креативность, организаторские навыки, Adobe Photoshop, техническое обслуживание, умение работать в условиях многозадачности, работа с базами данных, обучение и развитие, ориентация на результат
8	Управление проектами	Руководитель проектов, руководитель проекта, менеджер проектов, Project Manager, руководитель IT-проектов	управление проектами, Agile Project Management, английский язык, Scrum, работа в команде, Atlassian Jira, MS Project, организаторские навыки, управление командой, Project Management, ведение переговоров, MS PowerPoint, бизнес-анализ, деловая коммуникация, SQL, разработка технических заданий, аналитическое мышление, product management, управление продуктом, MS Visio
9	Разработка на 1С	Программист 1С, ведущий программист 1С, разработчик 1С, консультант 1С, программист 1С (удаленно)	1С программирование, 1С:Предприятие 8, 1С:Зарплата и управление персоналом, 1С:Бухгалтерия, 1С:Управление торговлей, 1С:Документооборот, ERP-системы на базе 1С, 1С:Управление производственным предприятием, 1С:Комплексная автоматизация, 1С:Управление предприятием, обновление конфигурации 1С, разработка технических заданий, MS SQL, постановка задач разработчикам, бухгалтерский учет, 1С:Предприятие, SQL, ERP, XML
10	Анализ данных	Аналитик, аналитик данных, Data Scientist, Data Engineer, Data Analyst	SQL, Python, MS SQL, анализ данных, аналитическое мышление, базы данных, MS PowerPoint, MS SQL server, работа с базами данных, английский язык, PostgreSQL, Power BI, Tableau, Oracle, MS Excel, математическая статистика, ETL, Oracle PL/SQL, Hadoop, СУБД
11	Интернет-реклама и графический дизайн	Интернет-маркетолог, графический дизайнер, технический писатель, SEO-специалист, бизнес-аналитик	Adobe Photoshop, Adobe Illustrator, Figma, креативность, UX, UI, Google Analytics, информационная безопасность, грамотная речь, грамотность, Яндекс.Метрика, AutoCAD, HTML, графический дизайн, интернет-реклама, контекстная реклама, MS PowerPoint, SQL, информационные технологии, пользователь ПК
12	Frontend-разработка	Frontend-разработчик, Frontend Developer, Frontend developer, Web-разработчик, Senior Frontend Developer	JavaScript, Git, CSS, HTML, React, HTML5, TypeScript, PHP, CSS3, Node.js, MySQL, jQuery, Redux, Vue.js, OOP, SQL, PostgreSQL, Linux, webpack, 1С-Битрикс
13	Тестирование	QA Engineer, QA Automation Engineer, тестировщик, QA engineer, тестировщик ПО	SQL, тестирование, Atlassian Jira, функциональное тестирование, Git, Java, QA, Python, Test Case, Linux, тестирование пользовательского интерфейса, API, Postman, Atlassian Confluence, регрессионное тестирование, английский язык, REST, XML, проведение тестирований, Android

Окончание табл. 8 / End of the table 8

Кластер	Название кластера	Наименования вакансий	Требуемые навыки
14	Продажи	Менеджер по продажам, менеджер по работе с клиентами, менеджер интернет-магазина, менеджер проектов, менеджер по продажам B2B	активные продажи, ведение переговоров, B2B продажи, поиск и привлечение клиентов, телефонные переговоры, грамотная речь, навыки продаж, работа в команде, проведение презентаций, развитие продаж, холодные продажи, CRM, деловое общение, прямые продажи, клиентоориентированность, навыки презентации, деловая коммуникация, деловая переписка, ориентация на результат, заключение договоров
15	Разработка мобильных приложений	IOS разработчик, Android разработчик, Python разработчик, Android Developer	Git, SQL, PostgreSQL, OOP, C#, Java, Python, Linux, PHP, MySQL, Kotlin, Docker, Android, iOS, .NET Framework, SWIFT, Redis, MS SQL, C++, golang

Таблица 9 / Table 9

**Характеристики кластеров вакансий****Characteristics of vacancy clusters**

Кластер	Название кластера	Средний минимум зарплаты	Средний максимум зарплаты	Требуемый опыт работы (мода)	Процент вакансий с требуемым опытом работы	График работы (мода)	Процент вакансий с данным графиком работы
1	DevOps-инженерия	160 000	227 000	От одного года до трех лет	47,5	Полный день	65,7
2	Системное администрирование	68 000	84 000	От одного года до трех лет	54,8	Полный день	85,8
3	Администрирование и техническая поддержка	125 000	184 000	От одного года до трех лет	45,0	Полный день	67,2
4	Java-разработка	210 000	302 000	От трех до шести лет	48,9	Полный день	56,6
5	Системный анализ и проектирование	131 000	175 000	От одного года до трех лет	56,4	Полный день	70,8
6	Администрирование проектов	67 000	95 000	От одного года до трех лет	53,7	Полный день	76,1
7	Техническая поддержка и сервис	68 000	95 000	От одного года до трех лет	55,0	Полный день	73,6
8	Управление проектами	135 000	175 000	От трех до шести лет	47,0	Полный день	79,0
9	Разработка на 1С	123 000	155 000	От трех до шести лет	48,4	Полный день	77,7
10	Анализ данных	123 000	173 000	От одного года до трех лет	56,2	Полный день	76,6
11	Интернет-реклама и графический дизайн	83 000	117 000	От одного года до трех лет	48,9	Полный день	75,9
12	Frontend-разработка	142 000	213 000	От одного года до трех лет	48,7	Полный день	51,0
13	Тестирование	131 000	175 000	От одного года до трех лет	63,7	Полный день	60,5
14	Продажи	77 000	145 000	От одного года до трех лет	59,1	Полный день	74,8
15	Разработка мобильных приложений	169 000	254 000	От одного года до трех лет	45,9	Полный день	54,3

## 4. Выводы

Предлагаемый подход к интеллектуальному анализу данных о вакансиях продемонстрирован на примере отрасли ИТ и объявлений о вакансиях в городах Москва и Санкт-Петербург. Предлагаемый подход позволяет выявить актуальные группы профессий и соответствующие им, востребованные работодателями навыки, умения и знания технологий. Анализ может проводиться в разрезе отраслей, регионов. Описание кластеров может быть дополнено суммаризацией должностных обязанностей, характерных для каждого кластера, на основе обработки текстовых описаний вакансий.

Рекомендуется выполнять анализ данных о вакансиях на регулярной основе не реже одного раза в год для мониторинга изменений на рынке труда и выявления новых профессиональных трендов. Результаты анализа данных о вакансиях могут быть использованы академическими руководителями образовательных программ в высших учебных заведениях для обновления учебных планов. Предложенный подход к анализу данных о вакансиях может быть реализован в рамках информационно-аналитической системы, предназначенной для поддержки решений руководства высших учебных заведений, а также в качестве информационно-аналитического сервиса для абитуриентов, выбирающих будущую профессию, и их родителей.

### Финансирование

Это исследование не получило какого-либо гранта от финансирующих агентств в государственном, коммерческом или некоммерческом секторах.

### Funding

This study did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

## Список источников / References

1. Дугарова Д. Ц., Старостина С. Е. Профессиональные стандарты как фактор гарантий качества подготовки выпускников образовательных организаций высшего и профессионального образования. *Педагогическое образование в России*. 2019;(3):5–14. DOI: 10.26170/по19-03-01 [Dugarova D. T., Starostina S. E. Professional standards as a quality guarantee factor training graduates of educational organizations higher and professional education. *Pedagogical Education in Russia*. 2019;(3):5–14. (In Russian.) DOI:10.26170/по19-03-01]
2. Omelchenko I., Antonova G., Danilina M., Popkov S., Botasheva L. Digitalization in the labor market in Russia. *SHS Web of Conferences*. 2021;(93):01021. DOI: 10.1051/shsconf/20219301021
3. Elias P., Purcell K. SOC (HE): A classification of occupations for studying the graduate labour market. Institute for Employment Research, University of Warwick, Coventry, UK: 2004. 46 p. Available at: <http://www2.warwick.ac.uk/fac/soc/ier/research/completed/7yrs2/rp6.pdf>
4. Litecky C. R., Arnett K. P., Prabhakar B. The paradox of soft skills versus technical skills in is hiring. *Journal of Computer Information Systems*. 2004;45(1):69–76. DOI: 10.1080/08874417.2004.11645818
5. Lee C. K., Han H. J. Analysis of skills requirement for entry-level programmer / analysts in Fortune 500 corporations. *Journal of Information Systems Education*. 2008;19(1):17–27. Available at: <https://www.thefreelibrary.com/Analysis+of+skills+requirement+for+entry-level+programmer%2Fanalysts+in...-a0233828125>
6. Stal J., Paliwoda-Pekosz G. Fostering development of soft skills in ICT curricula: A case of a transition economy. *Information Technology for Development*. 2019;25(2):250–274. DOI: 10.1080/02681102.2018.1454879
7. Stephany F. One size does not fit all: Constructing complementary digital reskilling strategies using online labour market data. *Big Data & Society*. 2021;8(1). DOI: 10.1177/20539517211003120
8. Zakaria N., Ismail N. N., Ana A. The relationship between employability skill and career choices among vocational skill students. *Innovation of Vocational Technology Education*. 2016;12(2):81–84. DOI: 10.17509/invotec.v12i2.6210
9. Jansen B. J., Jansen K. J., Spink A. Using the web to look for work: Implications for online job seeking and recruiting. *Internet Research*. 2005;15(1):49–66. DOI: 10.1108/10662240510577068
10. Boselli R., Cesarini M., Marrara S., Mercurio F., Mezzanzanica M., Pasi G., Viviani M. WoLMIS: A labor market intelligence system for classifying web job vacancies. *Journal of Intelligent Information Systems*. 2018;51(3):477–502. DOI: 10.1007/s10844-017-0488-x
11. Jackson M., Goldthorpe J. H., Mills C. Education, employers and class mobility. *Research in Social Stratification and Mobility*. 2005;(23):3–33. DOI: 10.1016/S0276-5624(05)23001-9
12. Jackson M. How far merit selection? Social stratification and the labour market. *British Journal of Sociology*. 2007;58(3):367–390. DOI: 10.1111/j.1468-4446.2007.00156.x
13. Dörfler L., van de Werfhorst H.G. Employers' Demand for Qualifications and Skills. *European Societies*. 2009;11(5):697–721. DOI: 10.1080/14616690802474374
14. Kureková L. M., Beblavý M., Haita C. Qualifications or Soft Skills? Studying Job Advertisements for Demand for Low-Skilled Staff in Slovakia. *NEUJOBS Working Paper D.4.3.3, 2012*. 36 p. DOI: 10.2139/ssrn.2402729
15. Kureková L. M., Beblavý M., Thum-Thysen A. Using online vacancies and web surveys to analyse the labour market: A methodological inquiry. *IZA Journal of Labor Economics*. 2015;(4):1–20. DOI: 10.1186/s40172-015-0034-4
16. Sasser Modestino A., Shoag D., Balance J. Upskilling: Do employers demand greater skill when workers are plentiful? *The Review of Economics and Statistics*. 2020;102(4):793–805. DOI: 10.1162/rest\_a\_00835
17. Beblavý M., Fabo B., Lenaerts K. Skills requirements for the 30 most-frequently advertised occupations in the United States: An analysis based on online vacancy data. *Centre for European Policy Studies (CEPS). Special report*. 2016. 41 p. Available at: <http://ssrn.com/abstract=2749549>
18. Carnevale A. P., Jayasundera T., Repnikov D. Understanding online job ads data: A technical report. *Georgetown University, McCourt School on Public Policy, Center on Education and the Workforce*. 2014. 20 p. Available at: [https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM\\_Tech\\_Web\\_.pdf](https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM_Tech_Web_.pdf)
19. Föll P., Thiesse F. Aligning is curriculum with industry skill expectations: A text mining approach. *Proceedings of the 25th European Conference on Information Systems (ECIS)*. 2017(6):2949–2959. Available at: <https://core.ac.uk/download/pdf/301372478.pdf>
20. Marrara S., Pasi G., Viviani M., Cesarini M., Mercurio F., Mezzanzanica M., Pappagallo M. A language modelling approach for discovering novel labour market occupations from the web. *Proceedings of the International conference on Web intelligence WI '17*, 2017:1026–1034. DOI: 10.1145/3106426.3109035
21. Poonnawat W., Pacharawongsakda E., Henchareonlert N. Jobs analysis for business intelligence skills requirements in the ASEAN region: A text mining study. *Advances in Intelligent Informatics, Smart Technology and Natural Language Processing. iSAI-NLP 2017. Advances in Intel-*

*ligent Systems and Computing*. 2019;(807):187–195. DOI: 10.1007/978-3-319-94703-7\_17

22. Pejić-Bach M., Bertoucel T., Meško M., Krstiić Ž. Text mining of industry 4.0 job advertisements. *International Journal of Information Management*. 2020;(50):416–431. DOI: 10.1016/j.ijinfomgt.2019.07.014

23. Ternikov A. A. Skill-based clustering algorithm for online job advertisements. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*. 2022;22(2):250–265. DOI: 10.18500/1816-9791-2022-22-2-250-265

#### **Информация об авторе**

**Морозова Юлия Александровна**, канд. экон. наук, доцент департамента бизнес-информатики, Высшая школа бизнеса,

Национальный исследовательский университет «Высшая школа экономики», г. Москва, Россия; *ORCID*: <https://orcid.org/0000-0002-2849-6218>; *e-mail*: yumorozova@hse.ru

#### **Information about the author**

**Yulia A. Morozova**, Candidate of Sciences (Economy), Associate Professor at the Department of Business Informatics, Graduate School of Business, National Research University Higher School of Economics, Moscow, Russia; *ORCID*: <https://orcid.org/0000-0002-2849-6218>; *e-mail*: yumorozova@hse.ru

**Поступила в редакцию / Received**: 22.08.2022.

**Поступила после рецензирования / Revised**: 25.09.2022.

**Принята к печати / Accepted**: 27.09.2022.