

## Обзор исследований рынка труда IT-специалистов, основанных на подходе анализа больших данных

Одной из наиболее динамично развивающихся отраслей в условиях перехода к цифровой экономике является отрасль информационных технологий (далее - ИТ), которая также стала важной частью инновационного развития государства. На данный момент, развитие большинства отраслей напрямую зависит от развития информационных технологий, технические специалисты с определенными наборами компетенций востребованы в областях, связанных с экономикой, управлением торговлей и т.д.

Недостаточное развитие собственного ИТ-рынка имеет негативное влияние на развитии прочих отраслей государства.

Кризисные ситуации в стране неизбежно отражаются на весь рынок труда России, в частности, и на ИТ-рынок. На ежемесячном обзоре рынка труда от веб-портала по поиску работы HeadHunter (далее - hh.ru) можно наблюдать изменения количества вакансий и количества резюме, начиная с января 2019 года по март 2023 [1]. На Рисунке 1 красным цветом представлен график изменения количества вакансий, синим цветом график изменения количества резюме относительно января 2019 года.

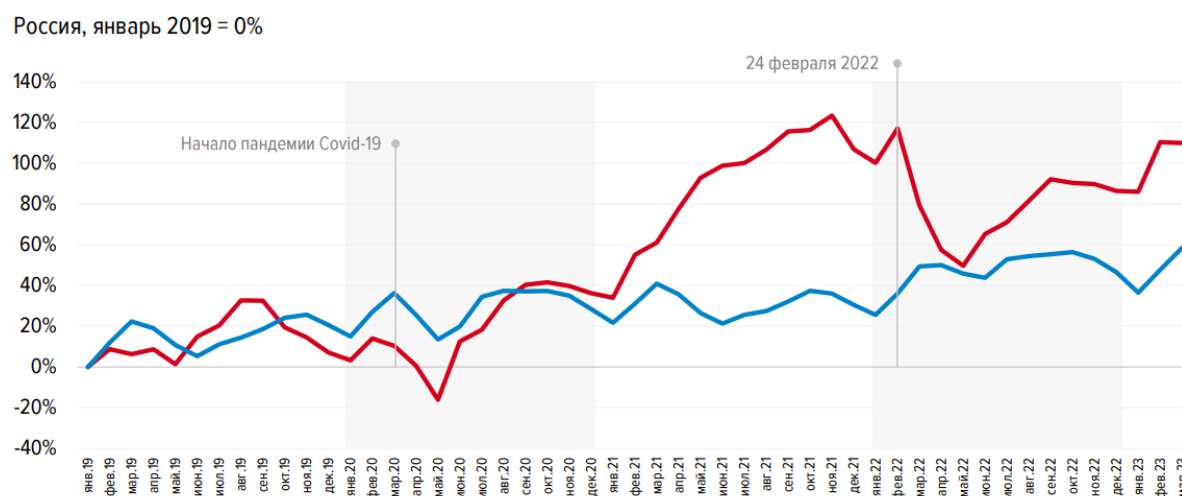


Рис 1. Динамика числа активных вакансий и резюме относительно января 2019

Большим толчком для ускоренной цифровизации экономики России послужила пандемия коронавируса, об этом также упомянул Роберт Фариш, вице-президент IDC(International Data Corporation) в России и СНГ, который поделился данными по итогам 2020 года. По данным IDC, в 2020 году рост ИТ-рынка в рублях составил 14% по отношению к 2019 году. Все это привело к значительному увеличению спроса на специалистов в ИТ-сфере. Следствием этого, увеличился поток молодых специалистов, которые, в большинстве своем, уступали по компетенциям более опытным кандидатам. И несмотря на то, что спрос на ИТ-специалистов не успевал за увеличением количества резюме, нехватка высококвалифицированных специалистов сохранялась.

Из-за множественных ограничений, наложенных на экономические возможности России, из страны ушло много зарубежных компаний, что привело к резкому снижению количества вакансий, а также. Данное обстоятельство отразилось на ИТ-рынке, множество квалифицированных специалистов или остались без работы, при этом сохраняя высокие зарплатные ожидания, или переехали в другие страны, что ощутимо сказывается на потребности в новых квалифицированных специалистах.

Таким образом, в сфере информационных технологий России имеется острая необходимость формирования кадров в сфере ИТ. Государство разработало федеральный проект «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации», которая реализуется в рамках перечня инициатив социально-экономического развития Российской Федерации до 2030 года [2]. Целью проекта является создание возможностей для формирования востребованных рынком труда цифровых компетенций.

Для наилучшего понимания необходимых компетенций для достижения квалифицированного уровня, имеет смысл изучить требования к специалистам на основе анализа вакансий, размещенных на крупнейших интернет-порталах поиска работы [3]. Самым крупным является сервис hh.ru, на котором, на момент написания статьи, опубликовано около ста тысяч активных вакансий в категории «Информационные технологии, системная интеграция, интернет».

Получив информацию из вакансий и проанализировав ее, можно составить перечень необходимых компетенций в зависимости от выбранных направлений, опыта работы, региона. Данная задача стала основной для коллектива авторов, решением которой коллектив занимается в настоящий момент.

Однако, прежде всего необходимо удостовериться в актуальности проводимого исследования и в наличии аналогичных исследований. Для этого было отобрано и проанализировано около 20 статей со схожими направлениями исследований. Для 5 наиболее подходящих статей были сформированы рецензии. Основными критериями при анализе статей являлись входные данные, их количество и методы очистки, а также подходы к кластеризации и классификации [4-5] полученных результатов.

Первая статья в рассматриваемом цикле была написана преподавателями различных университетов России. Работа опубликована в Британском научном журнале “Journal of Physics: Conference Series”[6] с H-индексом равным 85. В нем читателям представляют последние достижения в области физики, представленные на международных конференциях. Журнал является рецензируемым изданием открытого доступа издательства IOP Publishing. Данное издательство индексируется в таких системах, как: Inspec, Scopus, INSPIRE-HEP, MathSciNet, ISI Proceedings, Chemical Abstracts и множестве других.

В статье производится анализ рынка ИТ специальностей посредством построения карты компетенций разработчика, благодаря которой можно сделать выводы о наиболее востребованных навыках для специалистов различных отраслей.

Для достижения поставленной цели авторами была использована тестовая база данных вакансий, полученная с портала «Работа в России» - федеральной государственной информационной системы Федеральной службы по труду и занятости. База включала

более трех миллионов актуальных за период с октября 2015 года по июнь 2017 года вакансий, выполненных Центрами занятости, личными кабинетами работодателей на различных порталах и информационными системами работодателей. Ключевыми данными, внесшими свой вклад в исследование, авторами были выделены такие позиции как: требования, квалификация, дата добавления, заработная плата.

Перед началом работы с полученными данными, авторами была проведена их очистка, включающая в себя удаление HTML-тегов, дубликатов, опечаток, вакансий с отсутствующей квалификацией.

Для определения общих требований к специалистам из подготовленного набора данных были использованы методы кластеризации и семантического анализа. В ходе работы авторами был получен граф, в котором повторяющиеся фразы в критерии “категория” были определены как вершины графа, а отношения между фразами — как ребра графа. Для оценки правильности размещения кластеров был использован показатель модульности, благодаря которому была проведена оценка качества его структуры. Из-за сложности задачи оптимизации данного показателя авторами был выбран метод Лувена для решения этой проблемы, благодаря его высокой скорости и точности расчетов, а также возможности перерасчета модульности с другим параметром разрешения.

В ходе исследования авторами была получена карта компетенций специалистов.

Результаты проведенной работы предлагается использовать при разработке профессиональных стандартов и образовательных программ (как начального, так и среднего образования). Перспективными направлениями развития отмечаются изучение резюме в соответствующей отрасли для построения карт компетенций специалистов, присутствующих в настоящее время на рынке труда, а также изучение карт требуемых компетенций в контексте времени.

Статья имеет множество сильных сторон. Авторы визуализируют результаты своего исследования, что повышает понимание проведенного анализа, используют кластеризацию, опираются на множество различных источников. Однако, для достижения улучшения целевой метрики есть возможность изменить некоторые составляющие исследования. Например, увеличить выборку ключевых параметров, конечное число которых в ходе работы составило 26450. Также произвести более глубокую очистку данных, добавив контекстный анализ, лемматизацию, обработку шумовых значений и интеллектуальный анализ данных.

В статье «Анализ спроса на компетенции в области информационных технологий от лидирующих российских компаний»[7], опубликованной в журнале "E-Management.", который издается с 2018 года и включен в перечень ВАК по следующим направлениям: региональная и отраслевая экономика, мировая экономика, менеджмент, авторами являются кандидаты экономических наук Константинова Любовь и Крамаренко Инна, являющиеся доцентами кафедры математических методов в экономике и управлении в Государственном университете управления города Москва, изучают спрос на компетенции информационных технологий среди лидирующих российских компаний «Яндекс» и «VK Group».

Для исследования использовались данные из открытых источников, сбор данных осуществлялся следующим методом: исходный массив сформирован автоматически при помощи парсера ParseHub, на основе базы данных сайта по рекрутингу HeadHunter, а также официальных сайтов «Яндекс» и «VK Group», и скорректирован вручную. Массив сформирован из 256 вакансий. Также авторами была проведена группировка по самым востребованным ИТ-специальностям и общим формулировкам требований к этим специальностям, что позволило выявить ключевые навыки специалистов актуальных ИТ-направлений. Результаты исследования могут быть использованы для будущего более глубокого исследования потенциала рынка труда в сфере ИКТ, а также учтены при формировании компетенций специалистов, желающих работать в области ИТ в дальнейшем.

Актуальность статьи подтверждает активное развитие ИТ-отрасли, а также рост требований к компетенции ИТ-сотрудников, по данным HR-департамента Страхового Дома ВСК на конец 2021 г. количество вакансий в сфере ИТ увеличилось на 64%. Рассматриваются современные требования к ИТ-специалистам актуальных ИТ-направлений от «Яндекс» - разработка, аналитика, эксплуатация, от «VK Group» - серверная разработка, клиентская разработка, аналитика.

В исследовании были использованы совокупности теоретических и эмпирических методов, а также приведены алгоритм и результаты агрегирования групп ИТ-специалистов и их профессиональных компетенций, востребованных работодателями.

Преимуществом статьи являются ее актуальность на сегодняшний день, поскольку статья была опубликована в 2022 году. Также, данные взяты у крупных российских ИТ-компаний - «Яндекс» и «VK Group», которые активно развивают свои сервисы и внедряют новые технологии в них, поэтому вакансии и требования в них актуализированы на сегодняшний день. Однако, данный пункт также является и недостатком, поскольку уменьшается репрезентативность общей совокупности вакансий рынка ИТ, а также не рассматриваются иностранные компании. Кроме того, в статье не рассматривается динамика требований к ИТ-специалистам для анализа дальнейшего роста требований. С точки зрения возможного увеличения научной значимости статьи, необходимо увеличить количество компаний, в том числе рассмотреть и зарубежные, а также увеличить количество вакансий и провести анализ динамики требований, для выводов о дальнейших изменений в требованиях компетенции ИТ-специалистов.

Далее была рассмотрена публикация «Обзор рынка труда в ИТ-сфере в начале 2021 года в России и Санкт-Петербурге» [8], в которой рассматривается развитие ИТ-рынка на момент первого квартала 2021 года с учетом статистика с 2010 года. Аналитики из hh.ru выделили несколько трендов на основе 250 000 вакансий в сфере «ИТ, системная интеграция, интернет». Данная статья интересна тем, что информация получена на основе вакансий с сайта hh.ru, и проанализирована своими же специалистами из области анализа данных.

Статья является не научной, а скорее общеисследовательской. И, как следствие, в статье не раскрыты используемые методы для анализа и подготовки сырых данных, а также нет подробного описания источника данных.

Тренды, выделенные в статье, основываются на полученных статистикой фактах, что хорошо отражается в объективности статьи в целом. Однако, большая часть сравнений приводится на основе одного квартала или за предыдущий год, это позволяет предполагать изменения в недалеком будущем, но сложно сделать какой-либо прогноз на более длительные сроки.

На момент написания рецензии, статья теряет свою актуальность, так как за время прошедшее с выпуска статьи ИТ-рынок сильно изменился. Эту проблему можно было бы решить выпуская аналогичные статьи ежегодно.

Четвертая статья - “Разработка и апробация технологии анализа актуальных вакансий”[9], написана преподавателями и кандидатами технических наук Оренбургского Государственного Университета. Работа опубликована в общероссийском научно-педагогическом журнале «Высшее образование в России». Учредителями журнала являются Московский политехнический университет и Ассоциация технических университетов. В этом журнале публикуются теоретические статьи, а также результаты эмпирических и практико-ориентированных исследований, материалы конференций, научные рецензии. Издание входит в Перечень рецензируемых научных изданий ВАК РФ. Журнал индексируется в системах: Российский индекс научного цитирования, Академия Google, ErihPlus, Scopus, СОЦИОНЕТ, WorldCat.

В статье анализируются вакансии ИТ-специалистов на рынке труда, различные требования работодателей к программистам при трудоустройстве, методы отбора и группировки умений и навыков. При анализе требований работодателей, авторы статьи опираются на данные, взятые с портала поиска работы HeadHunter. Выборка вакансий производилась по г. Екатеринбургу за июль 2021 года. К сожалению, авторы не указали метод отбора требований с каждой вакансии, вследствие чего сложно оценить точность приведенных данных.

В статье используются методы кластеризации для формирования групп похожих вакансий и определения общих требований к специалистам. Для оценки склонности к кластеризации выбрана статистика Хопкинса. Результатом вычислений является коэффициент, указывающий общую предрасположенность имеющихся данных к объединению в кластеры. Для самого процесса кластеризации использовался метод k-средних, для оценки качества кластеризации применялась метрика силуэта. В результате проведенной работы авторами выявлены наиболее востребованные навыки и умения для каждой специальности. Стоит отметить, что в статье использовался эффективный метод кластеризации. Однако, метод k-средних чувствителен к шуму и мало эффективен при большом количестве невязанных между собой данных.

Статья имеет множество сильных сторон. Авторы визуализируют результаты каждого этапа исследования, что повышает понимание проведенного анализа, используют кластеризацию, опираются на множество различных источников. Однако, для более точного результата, корректно было бы изменить некоторые параметры. Например,

увеличить период анализа рынка труда (в статье продолжительность анализа ограничивается одним месяцем). Также, в статье произведена небольшая выборка вакансий – всего 974 позиции. Для большей точности можно увеличить количество вакансий.

Подводя итог рассмотрению публикаций можно увидеть, что основная точка роста исследований заключается во входных данных, зачастую, это малое количество отобранных вакансий. Помимо этого, в большинстве статей не описываются методы очистки данных. Также, методы кластеризации и классификации полученных результатов применяются только в единицах исследований. На основе вышесказанного можно сделать вывод, что данная тема освещена не в полной мере и требует дополнительного исследования.

За последние несколько лет рынок труда в России сильно изменился [10], большое влияние на это оказало развитие ИТ-отрасли. На данный момент, сфера информационных технологий является одной из самых развивающихся на рынке труда. Несмотря на это, ИТ-рынку все еще не хватает квалифицированных специалистов. Для того, чтобы успевать за изменяющимися тенденциями в ИТ, необходимо понимать требуемые компетенции выдвигаемые нанимающей стороной.

#### Источники:

- 1) Презентация Краткий обзор ситуации на рынке труда за март 2023. Режим доступа: <https://hhcdn.ru/icms/10283096.pdf> (дата обращения: 20.05.2023)
- 2) Федеральный проект Развитие кадрового потенциала ИТ-отрасли. Режим доступа: <https://digital.gov.ru/ru/activity/directions/1085/#section-departments> (дата обращения: 20.05.2023)
- 3) Ливинская В. А., Исследование актуальных потребностей рынка средствами современных методов анализа данных. Наука и образование: актуальные проблемы естествознания и экономики. - 2021. - С. 234-237.
- 4) P. Berkhin, "Survey of Clustering Data Mining Techniques," Accrue Software, 2002. Available from: <https://web.archive.org/web/20070117165907/http://citeseer.ist.psu.edu/berkhin02survey.html> (date of the application: 20.05.2023)
- 5) MacQueen, James. Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Vol. 1. No. 14. 1967.
- 6) Gintciak A., Rytik A., Terin D., et al Developing competency profiles of IT specialists based on semantic analysis of vacancies. Available from: [https://www.researchgate.net/publication/337736163\\_Developing\\_competency\\_profiles\\_of\\_IT\\_specialists\\_based\\_on\\_semantic\\_analysis\\_of\\_vacancies](https://www.researchgate.net/publication/337736163_Developing_competency_profiles_of_IT_specialists_based_on_semantic_analysis_of_vacancies) (date of the application: 20.05.2023)
- 7) Константинова Л. А., Крамаренко И. В., Анализ спроса на компетенции в области информационных технологий от лидирующих российских компаний. Режим доступа: <https://cyberleninka.ru/article/n/analiz-sprosa-na-kompetentsii-v-oblasti-informatsionnyh-tehnologiy-ot-lidiruyuschih-rossiyskih-kompaniy> (дата обращения: 20.05.2023)

- 8) Обзор рынка труда в ИТ-сфере в начале 2021 года в России и Санкт-Петербурге. Режим доступа: <https://rostov.hh.ru/article/28685> (дата обращения: 20.05.2023)
- 9) Шухман А. Е., Гришина Л. С., Легашев Л. В., Парфенов Д. И., Какие компетенции ждет региональный рынок ИТ-специалистов? Разработка и апробация технологии анализа актуальных вакансий. Режим доступа: <https://www.elibrary.ru/item.asp?id=49537516> (дата обращения: 20.05.2023)
- 10) Леманова П. В. Теория рынка труда. Функционирование механизмов рынка труда. Успехи современной науки. 2016. Т. 3, № 7. С. 26-28

**Сведения об авторах:**

*Кадомцев Максим Игоревич* - к.т.н, зав. каф. Медиа технологий, Донской Государственный Технический Университет, 2023

*Борисов Дмитрий Витальевич* - студент магистр, Донской Государственный Технический Университет, 2023

*Цеменко Олег Игоревич* - студент магистр, Донской Государственный Технический Университет, 2023

*Омельченко Андрей Олегович* - студент магистр, Донской Государственный Технический Университет, 2023

*Омельченко Сергей Евгеньевич* - студент магистр, Донской Государственный Технический Университет, 2023

*Данилов Даниил Вадимович* - студент магистр, Донской Государственный Технический Университет, 2023