



В. В. Малахов, Л. Г. Смышляева, А. Н. Мелентьева, А. О. Окороков

## Использование больших данных в практиках профориентации школьников на медицинскую профессию

**Введение.** Дефициты кадрового обеспечения здравоохранения негативно сказываются на доступности и качестве медицинской помощи и, соответственно, на качестве жизни и благополучии людей. Для устранения этих дефицитов предлагается использовать новые форматы и инструменты профессиональной ориентации школьников на медицинские профессии.

**Цель исследования:** выявить возможности использования больших данных для развития практик профориентации школьников в сферу медицинского труда.

**Материалы и методы.** Исследование было построено на выявлении паттерна личных и профессиональных интересов медицинского сообщества на основе данных цифрового следа медицинских работников, студентов первого и выпускного курса ФГБОУ ВО «Сибирский государственный медицинский университет» (далее – СибГМУ) [Российская Федерация]. Для проведения исследования была использована технология «Пиксель» от социальной сети «ВКонтакте». В исследовании приняли участие 947 человек, на основе личных данных которых был сформирован паттерн интересов медицинского сообщества.

Применялись методы теоретического и структурного анализа, метод предиктивной аналитики и имитационного моделирования на основе данных нейронных сетей и парсинг.

**Результаты исследования.** Выявлено 6 признаков сегментации целевой аудитории (медицинское сообщество), на основе маркетингового подхода найдены, проанализированы и дифференцированы (явные и неявные) идентификационные признаки, отражающие причастность пользователей социальной сети «ВКонтакте» к медицинскому сообществу. На основе технологии «Пиксель» обнаружено 3 подуровня из категории явного профессионального интереса пользователя и 184 характеристики, на основе которых они формируются. Используя базу социальной сети «ВКонтакте» обработано 79 млн страниц пользователей по идентификационным признакам, выявлено, что 79 тыс. человек соответствуют явному профессиональному интересу из категории «Медицина». Проведен поиск и анализ интересов студентов первого (937 человек) и выпускного курса СибГМУ, а также действующих специалистов (врачей с опытом работы более 5 лет), на основе которого, выявлен паттерн интересов из категории неявных идентификационных признаков, состоящий из 14 категорий.

**Заключение.** Результаты исследования показали, что использование больших данных и анализ цифрового следа медицинского сообщества является эффективным инструментом развития практик профориентационной работы со школьниками и их родителями. Выявление паттернов личных и профессиональных интересов медицинского сообщества может быть использовано для повышения эффективности профориентационной работы со школьниками в направлении мотивированного выбора ими сферы медицинского труда как карьерного вектора. Работа с большими данными, дополняя имеющиеся профориентационные подходы, выступает способом понимания предпочтений и интересов каждого школьника, что важно для обеспечения персонализированного сопровождения профессионального самоопределения.

**Ключевые слова:** медицинские профессии, профориентация школьников, персонализированное сопровождение профессионального самоопределения, большие данные, нейронные сети, предиктивная аналитика

### Ссылка для цитирования:

Малахов В. В., Смышляева Л. Г., Мелентьева А. Н., Окороков А. О. Использование больших данных в практиках профориентации школьников на медицинскую профессию // Перспективы науки и образования. 2023. № 6 (66). С. 516-531. doi: 10.32744/pse.2023.6.30



V. V. MALAKHOV, L. G. SMYSHLYAEVA, A. N. MELENTIEVA, A. O. OKOROKOV

## Use of big data in schoolchildren's career guidance practices for the medical profession

**Introduction.** The healthcare staffing deficiency affects negatively the availability and quality of medical aid and, consequently, the people's quality of life and well-being. To eliminate this deficiency, it is proposed to use new formats and tools for schoolchildren's professional orientation towards medical professions.

**Purpose of the study:** to identify the potential of using big data for the development of schoolchildren's career guidance practices in the field of medical labour.

**Materials and methods.** The research was grounded on the identification of the pattern of the medical community's personal and professional interests, as based on the data of the digital footprint of medical workers, first and final year students of Federal state-funded educational institution of higher education "Siberian State Medical University" (hereinafter SibSMU) [Russian Federation]. The technology "Pixel" borrowed from the social network "VKontakte" was used for the study. The study involved 947 persons whose personal data formed a basis for constructing a pattern of medical community interests.

The study involved the methods of theoretical and structural analysis, predictive analytics and simulation modelling based on neural network data, as well as parsing.

**Results.** A total of 6 attributes of the target audience segmentation (the medical community) were revealed; several identification features (explicit and implicit) reflecting the involvement of the social network "VKontakte" users in the medical community were found, analysed and differentiated on the basis of marketing approach. Using the "Pixel" technology, 3 sublevels reflecting the users' explicit professional-interest category and 184 characteristics underlying their formation were found. Using the base of the social network "VKontakte", a total of 79 million pages of users were processed according to the identification features. It was found that 79 thousand people matched the obvious professional interest reflecting the category "Medicine". The authors undertook a search and analysis of the interests of the first-year (937 people) and final-year students of SibSMU as well as of professionals (medical doctors with more than 5 years of service record), following which a pattern of interests pertaining to the category of implicit identifying characteristics that included 14 categories was revealed.

**Conclusion.** The research results showed that the use of big data and the analysis of the medical community's digital footprint is an effective tool for the development of career guidance practices for schoolchildren and their parents. The identification of patterns of the medical community's personal and professional interests can be used for the improvement of the efficiency of career guidance work with schoolchildren towards their motivated choice of the medical sphere as a career vector. The work with big data, complementing the existing career guidance approaches, is a due way to understand the preferences and interests of every schoolchild, which is important for achieving personalised support of professional self-identification.

**Keywords:** medical professions, schoolchildren's career guidance, personalised support of professional self-identification, big data, neural networks, predictive analytics

### For Reference:

Malakhov, V. V., Smyshlyayeva, L. G., Melentjeva, A. N., & Okorokov, A. O. (2023). Use of big data in schoolchildren's career guidance practices for the medical profession. *Perspektivy nauki i obrazovaniya – Perspectives of Science and Education*, 66 (6), 516-531. doi: 10.32744/pse.2023.6.30

## ВВЕДЕНИЕ

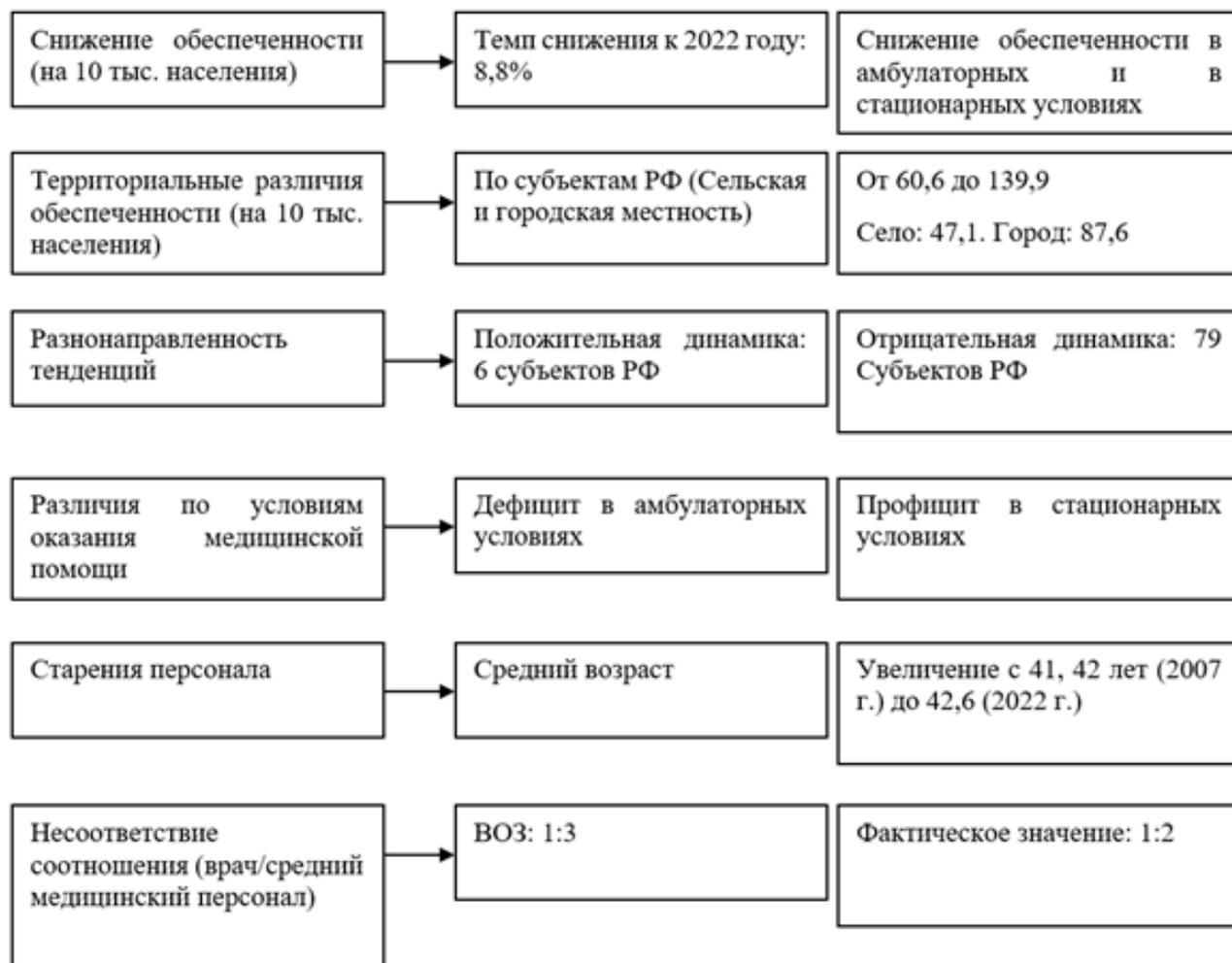
В последние два десятилетия международные организации, такие как Всемирная организация здравоохранения (далее – ВОЗ), ЮНЕСКО, Международная ассоциация университетов, активно занимаются проблемой кадрового обеспечения в здравоохранении. Одна из таких инициатив – «Стратегия развития здравоохранения до 2030 года», разработанная ВОЗ, которая призывает к обеспечению доступности и качества здравоохранения, включая обеспечение достаточного числа медицинского персонала [1]. Другим важным документом является «Рамочная конвенция по защите национальных меньшинств», принятая Советом Европы, которая призывает государства обеспечивать медицинское обслуживание, включая медицинский персонал, национальным меньшинствам и их членам [2]. Другим примером является Инициатива Международной ассоциации университетов «Глобальное партнерство по здравоохранению для устойчивого развития», которая призывает университеты разрабатывать и внедрять образовательные программы, направленные на подготовку медицинского персонала и улучшение доступности здравоохранения [3]. Наблюдается стабильное снижение обеспеченности врачами и специалистами со средним профессиональным медицинским образованием, что свидетельствует о необходимости принятия неотложных мер. За прошедший год, по оценкам ФГБУ «ЦНИИОИЗ» Минздрава России количество врачей в государственных медорганизациях России сократилось на 9,7 тысяч человек, а численность медицинского персонала продолжает уменьшаться. В 2022 году уровень обеспеченности средним медицинским персоналом уменьшился в 79 регионах. Средний показатель обеспеченности врачами составил 37,2 на 10 тысяч населения, а среднего медицинского персонала – 77,7 на 10 тысяч населения [4].

Таблица 1

## Обеспеченность медицинскими кадрами

	Врачи					Средний медицинский персонал				
	2020	2021	2022	Отклонение К 2020 г. в %	Отклонение К 2021 г. в %	2020	2021	2022	Отклонение К 2020 г. в %	Отклонение К 2021 г. в %
Российская Федерация	38	37,7	37,2	-2,1	-1,3	81,6	80,1	77,7	-4,8	-3
Северо-Западный федеральный округ	44,3	44,2	43,9	-0,9	-0,7	82	80,6	78,8	-3,9	-2,2
Южный федеральный округ	34,3	33,9	33,1	-3,5	-2,4	74,6	72,3	68,9	-7,6	-4,7
Северо-кавказский федеральный округ	35,8	35,6	35,7	-0,3	0,3	81,4	81	79,9	-1,8	-1,4
Приволжский федеральный округ	36,5	36,1	35,5	-1,6	-0,6	85,7	84,1	81,9	-4,4	-2,6
Уральский федеральный округ	34,7	34,4	34,2	-1,4	-0,6	88,8	87	84,6	-4,7	-2,8
Сибирский федеральный округ	37,8	37,3	36,6	-3,2	-1,9	86,7	84,8	82,1	-5,3	-3,2
Дальневосточный федеральный округ	40,1	39,6	38,9	-3	-1,8	84,2	83	80,7	-4,2	-2,8

Также отмечается ухудшение ситуации с кадрами в течение последних двух лет во всех федеральных округах, наиболее критическая ситуация наблюдается в Южном федеральном округе (-2,4%). Значительный дефицит врачебных кадров отмечен в 22 регионах, в том числе в Калужской области, Чукотском автономном округе, Чувашии, Тыве, Марий Эл, Калмыкии, Хакасии, Татарстане, Челябинской, Брянской, Липецкой, Кемеровской областях, Краснодарском крае и Севастополе. Расчеты проводились на основе данных форм федерального статистического наблюдения и территориального планирования, отчетных форм субъектов и территориальных программ госгарантий [5].



**Рисунок 1** Обеспеченность средним медперсоналом

Одним из способов решения этой проблемы является профессиональная ориентация школьников на медицинские профессии. Этот процесс является важной составляющей образовательной системы, направленной на помощь школьникам в выборе будущей профессиональной карьеры. Кроме того, профориентация может быть использована как инструмент для решения проблем кадрового обеспечения в медицинской сфере. Медицина требует высококвалифицированных кадров, недостаток которых может привести к ухудшению качества медицинской помощи и доступности. Кроме того, недостаточное количество медицинских кадров может препятствовать системе здравоохранения в эффективном реагировании на новые вызовы, такие как пандемия COVID-19 [6; 7]. Профессиональная ориентация школьников на медицинские профессии может быть ключевым решением этой проблемы.

Если больше учащихся заинтересовано в медицинских профессиях, то и кадровое обеспечение медицины будет более устойчивым. Однако, для того чтобы профориентация была эффективной, она должна быть современной и персонализированной [8]. На данный момент мы видим, что профориентационная система пользуется традиционными методами, в основе которых лежат: тесты на профессиональное самоопределение, опросы, профессиональные пробы, консультации и т.д. [9; 10]. Такие методы в современных реалиях недостаточно эффективны, с их помощью невозможно понять интересы человека в полной мере, а именно от этого зависит выбор профессии. Для того чтобы учитывать множество неструктурированной информации, такой как: интересы, привычки человека, его поведение, мировоззрение, круг общения и увлечения, необходимо собирать подобную информацию несколько лет (для более объективной оценки) и постоянно анализировать ее в динамике [11; 12]. Для этого необходимо использовать технологии больших данных. Во-первых, они позволяют анализировать огромные объемы данных, включающие информацию о медицинских специальностях, требованиях к квалификации и опыту работы, рыночных тенденциях и т.д. Это позволяет получить более точную и полную картину о ситуации на рынке труда и о том, какие специальности и направления в медицине востребованы в настоящее время и будут востребованы в будущем [13; 14]. Во-вторых, технологии больших данных могут помочь в создании персонализированных рекомендаций для школьников или студентов на основе их успеваемости, опыта работы, навыков и интересов. Это может помочь определить, какие медицинские профессии подходят конкретно им, что повысит мотивацию и шансы на успешную карьеру в выбранной сфере [15; 16]. В-третьих, анализ данных может помочь в прогнозировании тенденций на рынке труда в медицинской сфере, что может быть полезно как для тех, кто только начинают карьеру, так и для тех, кто уже работает в этой области и хочет планировать развитие карьеры на будущее. Современный инструментарий технологий больших данных, позволяет собирать, обрабатывать и анализировать не препарированные, разрозненные и неструктурированные массивы информации [17], с помощью data mining (нейронные сети, машинное обучение, искусственный интеллект) и строить из всего этого многообразия данных персонализированную траекторию дальнейшего развития и обучения человека, посредством имитационного моделирования и предиктивной аналитики данных [18; 19]. Такой системный подход к сбору и обработке информации о школьнике сможет дать объективное представление о его профессиональных интересах [20; 21]. В современном мире устроенность практик профориентационной работы начинает меняться. В ней все активнее начинают использоваться ИТ-технологии [22]. Но на данный момент это больше похоже на перенос традиционных методов на новые платформы. Вместе с тем, во многих сферах человеческой жизни технологии больших данных показали свою большую ресурсность. Особенно это касается области маркетинга [23]. Авторитетные эксперты отмечают, что их способы и методы можно использовать и для решения задач профориентации [24; 25]. Это можно сравнить с таргетированной и контекстной рекламой, которая предлагает человеку на основе интернет-запросов, телефонных разговоров, предпочтений, реальных перемещений, приложений, которые он использует, просмотров и поисков на видеохостингах, прослушивания музыки и подкастов, лайков и репостов в социальных сетях, товары и услуги, которые могут быть ему интересны [26; 27].

## МАТЕРИАЛЫ И МЕТОДЫ

Одним из самых базовых методов применения больших данных в маркетинге является сегментация целевой аудитории, так как небольшой группе людей с определенными общими признаками проще сделать более конкретное, адресное предложение. Также благодаря сегментации можно найти новую аудиторию и применить знание интересов целевой аудитории для подготовки и продвижения товаров или услуг. В ходе изучения маркетинговых методов были определены основные признаки сегментации аудитории. Приведем их характеристику в таблице 2.

Для решения задач исследования целесообразной стала сегментация целевой аудитории, что выступает одним из базовых методов при использовании больших данных в маркетинге. Это обусловлено тем, что небольшой группе людей с определенными общими признаками проще сделать более конкретное, адресное предложение. Кроме того, использование сегментации позволяет найти новую аудиторию и применить знание интересов целевой аудитории для подготовки и продвижения товаров или услуг. В ходе изучения маркетинговых методов были определены основные признаки сегментации аудитории. Приведем их характеристику в таблице 2.

**Таблица 2**

**Признаки сегментации целевой аудитории в маркетинге**

Признаки	Описание
Социально-демографические	К ним относятся пол, возраст, наличие семьи и события, происходящие с человеком на определенном этапе жизни, – окончание школы, университета, свадьба, рождение ребенка и другие. Для системы образования отдельного внимания заслуживает сегментирование аудитории по возрасту, так как для приема в университет или колледж необходимо разбивать аудиторию на возрастные составляющие и соответственно, что в свою очередь приведет к набору различных идентификационных признаков аудитории с которыми необходимо работать с учетом такой специфики.
Гиперлокальный таргетинг	Это один из видов сегментации аудитории по местоположению. Обращение в объявлении с привязкой к городу – популярный маркетинговый ход, но можно использовать и более узкую сегментацию – к тем точкам и объектам, которые люди посещают регулярно, бывают там или живут. Таких геопризнаков для сегментации много: места жительства: микрорайоны, кварталы, улицы, жилые комплексы, дома; культурные объекты: достопримечательности, известные места, площади, университеты, выставочные комплексы; места проведения досуга: парки, торговые центры, кафе, рестораны, развлекательные центры; спортивные объекты: фитнес-центры, спорткомплексы, стадионы; транспорт: железнодорожные станции, остановки, парковки; места работы: бизнес-центры, офисные помещения и др.
Профессиональные интересы	Этот тип интересов обеспечивает потребность пользователя в рекламируемом продукте. Например, продвижение одежды для медицинского персонала, целевая аудитория – врачи. Профессиональный интерес – медицина, и он тесно связан с рекламируемым продуктом. Чтобы поддерживать свой профессионализм, целевая аудитория может где-то учиться, читать литературу и блоги специалистов в интересующей области, а также покупать товары, которые помогут удовлетворить потребности в рамках профессионального интереса.
Парсер	Этот способ поиска позволяет обнаружить целевую аудиторию в скрытых сообществах социальных сетей – сообществах с нетипичными для продвигаемой тематики названиями.

Таргетинг на ключевые слова	Этот инструмент позволяет охватить пользователей со сформированным спросом – тех, кто ищет товары или услуги рекламодателя. Система при поиске аудитории учитывает поисковые запросы во «ВКонтакте», «Юле», «Одноклассниках», на главной Mail.ru и контентных проектах. Во внимание принимаются не только поиск, но и другие действия: переход в сообщества, просмотр товаров и т. п.
Личные интересы	Это бескрайнее поле холодной аудитории. Цель рекламы, которая охватывает таких пользователей – познакомить их с компанией и мотивировать задуматься над покупкой продукта. В образовании этот признак может дать огромное количество информации о потенциальном абитуриенте.

Описанные выше признаки сегментации целевой аудитории, по которым аккумулируется, анализируется и впоследствии используется информация в маркетинге можно адаптировать и решению задач системы высшего медицинского образования в контексте повышения эффективности работы с абитуриентами. Для построения модели такой работы можно взять за основу два признака сегментации аудитории – личные и профессиональные интересы. И, прежде всего, необходимо определить идентификационные признаки, характеризующие профессиональную группу «медицинские работники» (I этап), для того чтобы впоследствии использовать их для изучения и понимания абитуриентов медицинских вузов – школьников, профессионально нацеленных в сферу медицинского труда – при оценке и определении процентного соотношения схожести интересов (II этап).

Для того чтобы сформировать пул идентификационных признаков медицинского работника была проведена аналитическая работа (структурный анализ). Она была проведена посредством использования технологии «Пиксель» в социальной сети «ВКонтакте». Пиксель – это специальный фрагмент кода для отслеживания целевых действий пользователя на сайтах, в социальных и поисковых сетях. Технология, основой которой выступает нейронная сеть (часть технологий больших данных) позволяет фиксировать конверсии (любые действия пользователя) и моментально передавать эту информацию на платформу. Пиксель в основном используется в маркетинге для наиболее четкого сегментирования целевой аудитории. Отслеживание событий с помощью пикселя позволяет показывать рекламу наиболее заинтересованной аудитории – то есть, оптимизировать ее с учетом полученной информации. Для оптимизации модели машинного обучения (ML-модели) необходимы данные о том, что нужно учитывать при показе объявлений по выбранному целевому действию. Используя информацию, собранную пикселем, ML-модель в режиме реального времени подбирать наиболее релевантную аудиторию. Данные, которые получает пиксель с определенного сайта, зашифрованы, поэтому посмотреть или скачать их не получится – только использовать в продвижении. Российская социальная сеть была выбрана в качестве базы для исследования, так как обладает собственными сервисами по предоставлению услуг таргетинговой рекламы, это означает, что сеть самостоятельно собирает и обрабатывает все данные, с помощью элементов технологий больших данных (в данном случае используется технологии нейронных сетей) о пользователях для того чтобы составить портрет потребителя и сегментировать его.

Для данного исследования социальная сеть была весьма подходящей площадкой еще и потому, что дает возможность смотреть данные о пользователе по архивному запросу, где для участника доступен раздел реклама, в котором «ВКонтакте» собирает, анализирует и формирует интересы пользователя, исходя из его деятельности в сети. Сюда относятся: подписки на тематические сообщества, время прочтения постов,

процент дочитывания информации, реакция на посты (лайки, комментарии, игнорирование), добавление музыки в плейлист, поисковые запросы в интернете, отклики на рекламные посты. «ВКонтакте» делит все интересы пользователей на 2 группы: пользовательский интерес (полный анализ действий в социальной сети с алгоритмами, разработанными «ВКонтакте») и сторонний сегмент, схожий анализ только с алгоритмами разработанными другими компаниями. Все выше перечисленные особенности сбора данных и их последующая кластеризация относятся к методам обработки и сбору информации технологий больших данных, так как «ВКонтакте» использует машинное обучение, нейронные сети и технологии темных данных (автоматический сбор труднокластеризируемых данных), как часть технологий больших данных. Так, например, российская социальная сеть «ВКонтакте» сохраняет конфиденциальность онлайн-пользователей и одновременно помогает показывать рекламу на наиболее заинтересованную аудиторию. Технология позволяет автоматически определять все действия пользователя, среди которых: поиск по сайту; просмотр карточки или списка товаров; заполнение формы обратной связи; переход по ссылке в социальные сети или мессенджер; привязки телефонных номеров; время входа в приложение; история изменений имени; часто посещаемые места; информация о подписках на тематические сообщества; отметки «нравится»; фотографии; файлы; музыка; интересы; комментарии; репосты; какую информацию читали и на чем задерживали взгляд; хобби и ритм жизни; наличие определенных слов в переписке с другими людьми; историю посещений и данные об устройствах пользователей.

Если в течение 30 дней посетители совершали какое-либо из перечисленных действий, информация о нем отобразится в статистике, а само событие можно будет использовать для оптимизации работы с подобного рода аудиторией.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На первом этапе исследования необходимо было заранее разделить идентификационные признаки на явные (например, технология пиксель определит портрет целевой аудитории через личный или профессиональный признак по категории «Медицина») и неявные (например, признак личного интереса пользователя будет представлен категориями: «спорт», «музыка», «кино» и т.д.). После проведенного анализа был выявлен явный идентификационный признак в поле профессиональных интересов под названием «Медицина», то есть, социальная сеть «ВКонтакте», посредством собственной третирующей платформы уже сегментирует аудиторию сети по данной категории [28; 29]. Категория под названием «Медицина» – это слишком обширно, поэтому, на втором этапе опытно-экспериментальной работы был проведен анализ данных, из которых она формируется. С помощью оверлея ML-модели технологии «Пиксель» мы смогли увидеть траекторию формирования данной категории. Для наглядности определим, что категория «Медицина» – это уровень. Всё, чем будет данная категория представлена, это подуровни. С помощью анализа данных мы получили подуровни, которые показаны в таблице 3.

На третьем этапе исследовательской работы, используя базу пользователей социальной сети «ВКонтакте» (79 млн человек – пользователи сети из РФ), мы сделали выборку аудитории по идентификационным признакам, которые были найдены в результате второго этапа работы и выявили, что 79 тыс. человек в РФ (и 640 человек в Томской области) соответствуют явному профессиональному интересу из категории «Медицина».

Таблица 3

Подуровни категории явного профессионального интереса пользователя  
«Медицина»

Подуровень	Характеристики подуровня
Специальность	акушер; акушер-гинеколог; аллерголог; андролог; анестезиолог-реаниматолог; аритмолог; артролог; бактериолог; биохимик; вакцинолог; вегетолог; венеролог; вертебролог; вирусолог; врач; врач мрт; врач скорой помощи; врач функциональной диагностики; врач-цитолог; гастроэнтеролог; гематолог; гемостазиолог; генетик; гепатолог; гериатр (геронтолог); гинеколог; гинеколог-эндокринолог; гирудотерапевт; главврач (главный врач); гнатолог; гомеопат; дерматолог; дефектолог; диабетолог; диетолог; иммуногематолог; иммунолог; инфекционист; кардиолог; кардиохирург; кинезиолог; колопроктолог; комплементолог; косметолог; лазерный хирург; логопед; малоинвазивный хирург; маммолог; мануальный терапевт; массажист; медицинская сестра (медсестра); миколог; микробиолог; нарколог; невролог; нейропсихолог; нейрофизиолог; нейрохирург; неонатолог; нефролог; ожоговый хирург; окулист; онкогинеколог; онкодерматолог; онколог; онкопроктолог; онкоуролог; ортопед; остеопат; отоларинголог; отоневролог; офтальмохирург; паразитолог; педиатр; перинатолог; пластический хирург; подолог; проктолог; профпатолог; психиатр; психолог; психотерапевт; пульмонолог; радиолог; реабилитолог; реаниматолог; ревматолог; рентгенолог; репродуктолог; рефлексотерапевт; сомнолог; сосудистый хирург; спортивный врач; стоматолог; стоматолог-гигиенист; стоматолог-имплантолог; стоматолог-ортодонт; стоматолог-ортопед; стоматолог-пародонтолог; стоматолог-терапевт; стоматолог-хирург; сурдолог; терапевт; токсиколог; торакальный хирург; травматолог; трансплантолог; трансфузиолог; трихолог; УЗИ-специалист; уролог; фармаколог; фельдшер; физиатр; физиотерапевт; флеболог; фониатр; фтизиатр; химиотерапевт; хирург; челюстно-лицевой хирург; эмбриолог; эндокринолог; эндоскопист; эпилептолог;
Виды медицины	лечебная физкультура; патологическая анатомия; психология; судебная медицина; массаж; терапия; вирусология; патология; хирургия; ядерная медицина; анатомия; гигиена; психиатрия; психотерапия; эпидемиология; микробиология; нейропсихология; гомеопатия; онкология; гинекология; косметология; психофизиология; иммунология; травматология; акушерство; стоматология; физиотерапия; фармакология; кардиология; психопатология; педиатрия; ортопедия; неврология; гистология; трансплантология; гастроэнтерология; соматология; ринология; анестезиология; трансплантология; онкоцитология; гемеллология; гидropатия; паразитология; кардиохирургия; микология; наномедицина; дерматология; биомедицина; экомедицина; лекарство; лекарственный препарат; аптека; центр медицины; клиника китайской медицины; медицина институт; клиника эстетической медицины; ветеринарная медицина
Прочее	образование; здоровье; состояние здоровья; состояние организма; здравоохранение; история медицины; математика; химия; биология; естественные науки; человек; общество; помощь; школа медицины; медицинский университет; профессия; обучение; книги; работа; медик; языки; латынь; английский язык

На четвертом этапе исследования было осуществлено соотнесение полученных результатов с реальными (не обезличенными) данными медицинских работников, а также студентов первого и выпускного курса СибГМУ.

Были собраны и проанализированы данные семи выпускников СибГМУ, среди которых найдены дополнительные паттерны интересов, которые показаны в таблице 4.

Таблица 4

## Паттерны интересов выпускников СибГМУ

Категории	Процент совпадения категорий среди выпускников СибГМУ (%)
Здоровье	100
Кулинария	100
Спорт, спортивное питание	100

Продукты питания	100
Музыка	100
Отдых и путешествия	86
Фармацевтика	72
Медицинские учреждения	72
Медицина общая	72
Аптеки и лекарства	57
Красота и мода	57
Бизнес	57
Клиники	43
Детское здоровье	43
Диета и здоровое питание	43
Стоматология	29
Детское питание	29
Ветеринарные клиники	29
Аллергия, астма	29
Оптика	29
Общество	29

Далее был проведен идентичный анализ данных, но уже у трех специалистов (профессии) с многолетним стажем работы (данные в таблице 6).

Таблица 5

## Паттерны интересов состоявшихся специалистов

Категории	Процент совпадения категорий (%)
Здоровье	100
Спорт, спортивное питание	100
Продукты питания	100
Музыка	100
Отдых и путешествия	86
Медицина Фармацевтика	72
Медицинские учреждения	72
Медицина общая	72
Аптеки и лекарства	57
Красота и мода	57
Кулинария	57
Бизнес	57
Клиники	43
Общество	29

После сбора и анализа данных, представленных выше, было осуществлено сравнение паттернов интересов выпускников и состоявшихся специалистов. Это позволило найти несколько пересечений категорий (личных интересов), что говорит о том, что у медицинского сообщества существует определенный паттерн личных интересов (см. табл. 5).

Таблица 6

## Паттерны интересов медицинского сообщества

Категории	Процент совпадения категорий (%)
Медицина	100
Здоровье	100
Кулинария	100
Спорт, спортивное питание	100
Продукты питания	100
Музыка	100
Отдых и путешествия	90
Фармацевтика	70
Медицинские учреждения	70
Медицина общая	70
Аптеки и лекарства	60
Красота и мода	60
Бизнес	60
Общество	29

Таким образом, можно определить, что кроме явного профессионального интереса (категория медицина), медицинское сообщество интересуется как минимум 12 неявными категориями личных интересов среди которых: здоровье, кулинария, спорт, спортивное питание, продукты питания, музыка, отдых и путешествия, медицина (фармацевтика), медицинские учреждения, медицина общая, аптеки и лекарства, красота и мода, бизнес.

После сформированного паттерна личных интересов медицинского сообщества нами был проведен сравнительный анализ выявленных паттернов интересов на студентов первого курса СибГМУ и паттернов представителей профессионального медицинского сообщества. Результаты данного анализа приведены в таблице 7.

На данном этапе исследования было проанализировано 937 профилей студентов первого курса в социальной сети «ВКонтакте» с помощью метода парсинга из открытых данных.

Таблица 7

## Соответствие паттернов интересов студентов первого курса СибГМУ и паттернов представителей профессионального медицинского сообщества

Категории	Процент совпадения категорий (%)
Медицина	81
Здоровье	78
Кулинария	73
Спорт, спортивное питание	70
Продукты питания	68
Музыка	68
Отдых и путешествия	60
Медицинские учреждения	57

Медицина общая	57
Аптеки и лекарства	50
Красота и мода	50
Бизнес	43

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Проведенное исследование позволило расширить педагогические представления о способах и средствах развития практик профориентации школьников в сферу медицины. Технологии больших данных выступают одним из таких средств. Их использование позволяет выявить особенности (специфические признаки – паттерны) интересов людей, причастных к профессиональной медицинской деятельности. Анализ данных больших объемов и использование современных методов позволили выявить как явные, так и неявные интересы представителей современного медицинского сообщества, включая категории: здоровье, кулинария, спорт, музыка, отдых и путешествия, медицинские учреждения, бизнес и другие. Эти результаты представляют практически важную информацию для профориентационных программ. Доказательно установленные паттерны интересов медицинских работников можно использовать как диагностическую матрицу для выявления тех школьников, чьи интересы и социально-когнитивные предпочтения отражают совпадения с характеристиками этих паттернов, через анализ цифрового следа в социальных сетях и других цифровых пространствах. Такие школьники могут рассматриваться как потенциальные абитуриенты медицинских вузов, что определяет целесообразность более дифференцированной и целенаправленной профориентационной работы с ними и их родителями. Сравнительно-аналитический механизм, основанный на использовании больших данных при анализе интересов людей, причастных к профессиональной медицинской сфере, выступает новым дополнительным ресурсом в педагогическом сопровождении профориентационной работы со школьниками и их родителями. Данный механизм отражает целенаправленное педагогическое обращение к закономерностям киберсоциализации для содействия выпускникам школ в более осознанном профессиональном самоопределении и мотивированно-осмысленном построении будущей карьеры в отрасли здравоохранения. Кроме того, подход, основанный на сравнении паттернов интересов опытных специалистов в сфере здравоохранения и выпускников медицинских университетов, способствует лучшему пониманию последними своих карьерных потребностей в этой области – при условии педагогического сопровождения рефлексивного осмысления будущими врачами информации, полученной на основе анализа больших данных. Это открывает новые перспективы для обновления и повышения эффективности подходов к профориентационной работе, основанных на учете не только явных, но и неочевидных интересов студентов и школьников, содействуя им в осмысленном построении карьеры в сфере медицины.

Мы согласны, как с российскими исследователями (Е.С. Масленниковой, Н.В. Алексеевой и Л.В. Быковой [30]), так и с зарубежными учеными (E. Deci et al. [31]) в том, что профориентация школьников может быть более эффективной и прицельной с помощью анализа интересов и мотиваций медицинского сообщества. Это позволит более точно выявлять наиболее мотивированных будущих медицинских работников (абитуриентов медицинских вузов) и, соответственно, повышать качество медицинского образования и качество медицинской помощи.

Применение современных методов анализа данных, таких как ML-модели и парсинг открытых данных, в исследованиях современных проблем профессиональной ориентации школьников (в частности, на медицинскую профессию) представляется актуальным и обоснованным. Полученные нами данные согласуются с мнением авторов R. Lent и D. Brown о том, что такие методы позволяют более глубоко и детально исследовать интересы и предпочтения школьников и студентов, способствуя их правильному профессиональному самоопределению [32]. Результаты исследования позволяют установить общность интересов и социально-когнитивных предпочтений между выпускниками медицинского университета и уже работающими специалистами сферы здравоохранения, а обнаруженное пересечение интересов указывает на сходство между профессиональными ожиданиями и предпочтениями студентов первого курса и опытными специалистами. Использование метода парсинга данных из открытых источников, таких как социальные сети, позволило получить значительный объем информации о профилях будущих врачей в социальной сети, связанной с особенностями их профессиональной социализации. Это может быть использовано в работе со студентами-медиками в плане содействия им построения индивидуальных образовательных и карьерных стратегий. Представленный организационно-педагогический подход демонстрирует эффективность применения технологий больших данных и их потенциал в научных исследованиях, связанных с профессиональным самоопределением человека и развитием практик профессиональной ориентации школьников. Он создает новые возможности для изучения профессиональных интересов и предпочтений в различных сферах труда человека и обеспечивает получение более доказательных выводов на основе объективных данных, которые могут быть использованы педагогами, школьниками и их родителями.

---

## ЗАКЛЮЧЕНИЕ

Исследование представляет возможности применения современных методов анализа данных, таких как ML-модели и парсинг открытых данных, для повышения эффективности педагогического сопровождения профессионального самоопределения школьников в контексте обеспечения более точного понимания ими паттернов интересов, как одного из идентификационных признаков представителей профессиональных сообществ (в частности, медицинского).

На основе проведенного научно-педагогического исследования были получены результаты, которые позволяют обогатить современные практики профориентационной работы со школьниками в сфере здравоохранения за счет использования технологий больших данных. Для педагогического содействия школьникам в профессиональном самоопределении представляется важным обозначение им характеристик личностно-профессиональной идентичности представителей современного профессионального медицинского сообщества, чему также способствует, в частности, понимание профессиональных интересов и предпочтений студентов-будущих медиков и специалистов в сфере здравоохранения. Основой их выявления выступает обнаружение идентификационных признаков, таких, как явные и неявные интересы, сегментацию аудитории и сопоставление данных с реальными профессиональными и личными интересами. В результате исследования были выделены различные категории профессиональных интересов, связанных с медициной, а также установлены неявные категории личных

интересов медицинского сообщества. Обнаружено пересечение интересов между выпускниками медицинского университета и уже работающими специалистами. Эти результаты указывают на наличие определенных паттернов интересов в медицинском сообществе. Исследование позволило установить, что помимо явного профессионального интереса в медицине, медицинское сообщество обладает схожими интересами (идентификационными признаками) в области здоровья, кулинарии, спорта, музыки, отдыха и путешествий, медицинских учреждений, бизнеса и других сфер.

Данные результаты имеют практическую значимость для профориентационной работы среди будущих абитуриентов медицинских университетов, позволяя лучше понять их предпочтения и интересы, что важно для обеспечения персонализированного сопровождения профессионального самоопределения и планирования карьеры в сфере медицины.

В целом, результаты этого исследования могут быть полезными для разработки более эффективных программ профориентации школьников в сферу медицинского труда и карьерного планирования для абитуриентов и студентов медицинских вузов, а понимание и учет профессиональных и личных интересов помогут абитуриентам принимать эффективные решения о своей будущей карьере в медицинской сфере.

Дальнейшие исследования по проблеме исследования, представленной в статье, могут сосредоточиться на сравнении полученных результатов с данными других университетов (регионов), чтобы выявить общие тенденции или различия в профессиональных интересах абитуриентов, студентов и специалистов медицинской сферы.

## ЛИТЕРАТУРА

1. Федеральная служба государственной статистики // Здравоохранение в России. Режим доступа: [https://gks.ru/bgd/regl/b21\\_34/Main.htm](https://gks.ru/bgd/regl/b21_34/Main.htm) (дата обращения: 23.05.2023)
2. Шаронова А. П. Профориентация в медицинском образовании: теория и практика // Высшее образование сегодня. 2019. № 5-2. С. 100-114.
3. Кондрашова Л. В., Ковалев В. Н. Большие данные в образовании: анализ и возможности применения // Высшее образование сегодня. 2017. № 7-8. С. 44-52.
4. Поташник М. В. Применение больших данных в профессиональной ориентации студентов // Современное образование: проблемы и перспективы. 2017. № 3. С. 26-31.
5. Петрова И. В. Профориентация молодежи в медицинском образовании: проблемы и перспективы. // НИЦ «Радиотехника». 2016. № 5. С. 37-38.
6. Благов Ю. А., Панкова О. А. Применение методов анализа больших данных в процессе профессиональной ориентации // Информационные технологии в образовании. 2021. № 27. С. 67-73.
7. Коржевская Е. В. Профориентация в системе медицинского образования. // Высшее образование сегодня. 2022. № 3. С. 37-39.
8. Богомолова И. В., Ростовская Т. К. Профориентация студентов в медицинском колледже: методика, практика, проблемы // РостГМУ. 2020. № 2. С. 115-128.
9. Иванова А. А. Профессиональная ориентация студентов в медицинском образовании // Вестник СПбГМУ. 2017. № 5. С. 102-117.
10. Лаптева Н. А., Лаптева И. Н. Профориентация в медицинском образовании: теория и практика. // Высшее образование сегодня. 2018. № 4. С. 152-167.
11. Barron B. Interest and self-sustained learning as catalysts of development // A learning ecology perspective. Human Development. 2019. № 4. P. 193-224.
12. Hidi S., Renninger K. The four-phase model of interest development // Educational Psychologist. 2020. № 2. P.111-127.
13. Linnenbrink-Garcia L., Pekrun R. Students emotions and academic engagement: Introduction to the special issue // Contemporary Educational Psychology. 2020. № 3. P.223-237.
14. Savickas L., Nota L., Rossier J., Dauwalder J., Duarte E., Guichard J., Van Vianen E. Life designing: A paradigm for career construction in the 21st century // Journal of Vocational Behavior. 2019. № 3.P. 239-250.
15. Давыдова Е.А., Калинин А.А., Карабанова О.А. Профессиональная ориентация студентов медицинских вузов: проблемы и перспективы // Инновации в медицинском образовании. 2019. № 4. С. 11-16.
16. Григорьева Н.В. Профориентационная работа в медицинском колледже: теория и практика // Педагогическое образование в России. 2017. № 2. С. 73-80.

17. Кузнецов А.Н., Рыбин А.Ю., Кривоносов А.В. Профессиональная ориентация студентов-медиков: методы и результаты // *Профессиональное образование в современном мире*. 2018. № 5. С. 92-97.
18. Смирнова Т.В., Тарасова И.В. Профориентационная работа в медицинском вузе: состояние и перспективы // *Современные проблемы науки и образования*. 2016. № 1. С. 168-174.
19. Потапова Л.А., Коновалова Ю.В., Колосова О.В. Профессиональная ориентация студентов медицинского вуза: особенности и эффективность // *Наука и образование: современные тренды*. 2020. № 4. С. 97-102.
20. Зайцева Е.Ю., Никулина Е.А. Профориентационная работа в медицинском вузе: проблемы и перспективы // *Педагогика и психология профессионального образования*. 2018. № 2. С. 75-81.
21. Deci L., Ryan M. Self-determination theory: A macrotheory of human motivation, development, and health // *Canadian Psychology*. 2019. № 5. P. 182-185.
22. Tracey J., Robbins B. The interest-major congruence and college success relation: A longitudinal study // *Journal of Vocational Behavior*. 2021. № 3. P. 457-470.
23. Vallerand J., Houffort N. Passion at work: Toward a new conceptualization. // *Information Age Publishing*. 2022. № 2. P. 347-362.
24. Wigfield A., Eccles S. Expectancy-value theory of achievement motivation. // *Contemporary Educational Psychology*. 2022. № 3. P. 68-81.
25. Hartung J., Subich M. Beyond skills and interests: Career research and counseling in the postmodern era // *Journal of Career Development*. 2019. № 3. P. 87-96.
26. Hirschi A., Läge D. Career adaptability profiles and their relationship to adaptivity and adapting // *Journal of Vocational Behavior*. 2018. № 5. P. 249-262.
27. Creed A., Patton W., Prideaux A. Predicting change over time in career planning and career exploration for high school students // *Journal of Adolescence*. 2019. № 3. P. 19-29.
28. Nauta M., Epperson L., Kahn H. Applications of the social cognitive career theory to career counseling: Implications for enhancing collaboration between counselors and teachers // *Journal of Career Development*. 2020. № 2. P. 438-455.
29. Richardson S., Benner A. Longitudinal associations between early childhood temperament, interpersonal relationships, and career adaptability // *Journal of Vocational Behavior*. 2019. № 5. P. 1-11.
30. Масленникова Е.С., Алексеева Н.В., Быкова Л.В. Профориентация медицинских студентов: состояние и перспективы // *Педагогика и психология профессионального образования*. 2019. № 2. С. 60-65.
31. Gati I., Krausz M., Osipow H. A taxonomy of difficulties in career decision making // *Journal of Counseling Psychology*. 2019. № 4. P. 510-526.
32. Lent R., Brown, S., Hackett G. Toward a unifying social cognitive theory of career and academic interest, choice, and performance // *Journal of Vocational Behavior*. 2019. № 5. P. 79-122.

## REFERENCES

1. Federal State Statistics Service. Healthcare in Russia. Available at: [https://gks.ru/bgd/regl/b21\\_34/Main.htm](https://gks.ru/bgd/regl/b21_34/Main.htm) (accessed 23.05.2023)
2. Sharonova A. P. Career guidance in medical education: theory and practice. *Higher education today*, 2019, no. 5-2, pp. 100-114. (in Russ.)
3. Kondrashova L. V., Kovalev V. N. Big data in education: analysis and application possibilities. *Higher education today*, 2017, no. 7-8, pp. 44-52. (in Russ.)
4. Potashnik M. V. The use of big data in the professional orientation of students. *Modern education: problems and prospects*, 2017, no. 3, pp. 26-31. (in Russ.)
5. Petrova I. V. Career guidance of youth in medical education: problems and prospects. SIC "Radio Engineering", 2016, no. 5, pp. 37-38. (in Russ.)
6. Blagov Yu. A., Pankova O. A. Application of big data analysis methods in the process of professional orientation. *Information technologies in education*, 2021, no. 27, pp. 67-73. (in Russ.)
7. Korzhevskaya E. V. Career guidance in the system of medical education. *Higher education today*, 2022, no. 3, pp. 37-39. (in Russ.)
8. Bogomolova I. V., Rostovskaya T. K. Vocational guidance of students in medical college: methodology, practice, problems. *RostSMU*, 2020, no. 2, pp. 115-128. (in Russ.)
9. Ivanova A. A. Professional orientation of students in medical education. *Bulletin of SPbSMU*, 2017, no. 5, pp. 102-117. (in Russ.)
10. Lapteva N. A., Lapteva I. N. Career guidance in medical education: theory and practice. *Higher education today*, 2018, no. 4, pp. 152-167. (in Russ.)
11. Barron B. Interest and independent learning as catalysts for development. *The perspective of the ecology of learning. Human development*, 2019, no. 4, pp. 193-224.
12. Hidi S., Renninger K. Four-phase model of interest development. *Psychologist-pedagogue*, 2020, no. 2, pp.111-127.
13. Linnenbrink-Garcia L., Pekrun R. Emotions of students and academic involvement: introduction to a special issue. *Modern psychology of education*, 2020, no. 3, pp.223-237.
14. Savickas L., Nota L., Rossier J., Dauvalder J., Duarte E., Guichard J., Van Vianen E. Designing life: the paradigm of career building in the 21st century. *Journal of Professional Behavior*, 2019, no. 3, pp. 239-250.
15. Davydova E.A., Kalinin A.A., Karabanova O.A. Professional orientation of medical university students: problems and prospects. *Innovations in medical education*, 2019, no. 4, pp. 11-16. (in Russ.)

16. Grigorieva N.V. Career guidance work in medical college: theory and practice. *Pedagogical education in Russia*, 2017, no. 2, pp. 73-80. (in Russ.)
17. Kuznetsov A.N., Rybin A.Yu., Krivonosov A.V. Professional orientation of medical students: methods and results. *Vocational education in the modern world*, 2018, no. 5, pp. 92-97. (in Russ.)
18. Smirnova T.V., Tarasova I.V. Career guidance work in a medical university: state and prospects. *Modern problems of science and education*, 2016, no. 1, pp. 168-174. (in Russ.)
19. Potapova L.A., Konovalova Yu.V., Kolosova O.V. Professional orientation of medical university students: features and effectiveness. *Science and education: modern trends*, 2020, no. 4, pp. 97-102. (in Russ.)
20. Zaitseva E.Yu., Nikulina E.A. Career guidance work in a medical university: problems and prospects. *Pedagogy and psychology of vocational education*, 2018, no. 2, pp. 75-81. (in Russ.)
21. Deci L., Ryan M. Theory of self-determination: macrotheory of human motivation, development and health. *Canadian Psychology*, 2019, no. 5, pp. 182-185.
22. Tracy J., Robbins B. The relationship between interests and success in college: a longitudinal study. *Journal of Professional Behavior*, 2021, no. 3, pp. 457-470. (in Russ.)
23. Vallerand J., Holfort N. Passion at work: towards a new conceptualization. *Information Age Publishing House*, 2022, no. 2, pp. 347-362.
24. Wigfield A., Eccles S. The expected-value theory of achievement motivation. *Modern psychology of education*, 2022, no. 3, pp. 68-81.
25. Hartung J., Subich M. In addition to skills and interests: Career research and consulting in the postmodern era. *Journal of Career Development*, 2019, no. 3, pp. 87-96.
26. Hirshi A., Lage D. Profiles of career adaptability and their relationship with adaptability and adaptability. *Journal of Professional Behavior*, 2018, no. 5, pp. 249-262.
27. Creed A., Patton W., Prideaux A. Forecasting changes over time in career planning and career path research of high school students. *Journal of Adolescence*, 2019, no. 3, pp. 19-29.
28. Nauta M., Epperson L., Kan H. The application of the socio-cognitive theory of career to career guidance: implications for the expansion of cooperation between consultants and teachers. *Journal of Career Development*, 2020, no. 2, pp. 438-455.
29. Richardson S., Benner A. Longitudinal associations between early childhood temperament, interpersonal relationships and adaptability to career. *Journal of Professional Behavior*, 2019, no. 5, pp. 1-11.
30. Maslennikova E.S., Alekseeva N.V., Bykova L.V. Career guidance of medical students: state and prospects. *Pedagogy and psychology of vocational education*, 2019, no. 2, pp. 60-65. (in Russ.)
31. Gati I., Kraush M., Osipov H. Taxonomy of difficulties in making career decisions. *Journal of Counseling Psychology*, 2019, no. 4, pp. 510-526.
32. Lent R., Brown S., Hackett G. On the way to a unifying socio-cognitive theory of career and academic interest, choice and performance. *Journal of Professional Behavior*, 2019, no. 5, pp. 79-122.

**Информация об авторах**  
**Малахов Владислав Валерьевич**

(Россия г. Томск)

Аспирант Института развития педагогического образования  
ФГБОУ ВО «Томский государственный педагогический университет»  
Лаборант, Лаборатория развития образования  
ФГБОУ ВО СибГМУ Минздрава России  
E-mail: malakhov14@yandex.ru

**Смышляева Лариса Германовна**

(Россия г.Томск)

Доктор педагогических наук, доцент  
Заведующий лаборатории, Лаборатория развития образования  
ФГБОУ ВО СибГМУ Минздрава России  
E-mail: laris.s@mail.ru

**Мелентьева Александра Николаевна**

(Россия г.Томск)

Кандидат фармацевтических наук  
Начальник управления нового набора студентов  
ФГБОУ ВО СибГМУ Минздрава России  
E-mail: melesokol@gmail.com

**Окорок Александр Олегович**

(Россия г.Томск)

Проректор по учебной работе  
ФГБОУ ВО СибГМУ Минздрава России  
E-mail: prorektor.ur@ssmu.ru

**Information about the authors**

**Vladislav V. Malakhov**  
**(Tomsk, Russia)**

Postgraduate student of the Institute of Pedagogical Education Development  
Tomsk State Pedagogical University  
Laboratorian, Laboratory of Education Development  
Siberian State Medical University  
E-mail: malakhov14@yandex.ru

**Larisa G. Smyshlyayeva**

(Tomsk, Russia)

Dr. Sci. (Educ.), Associate Professor  
Head of Laboratory, Laboratory of Education Development  
Siberian State Medical University  
E-mail: laris.s@mail.ru

**Alexandra N. Melentieva**

(Tomsk, Russia)

Cand. Sci. (Pharm.),

Head of New Student Enrollment Department  
Siberian State Medical University  
E-mail: melesokol@gmail.com

**Alexander O. Okorokov**

(Tomsk, Russia)

Vice-rector for academic work  
Siberian State Medical University  
E-mail: prorektor.ur@ssmu.ru