

doi: 10.17586/2226-1494-2023-23-3-564-574

УДК 51-77: 519.25

Исследование тенденций взаимосвязи между профориентационными предпочтениями пользователей и их цифровыми следами в социальной сети

Анастасия Олеговна Хлобыстова¹, Максим Викторович Абрамов²✉,
Валерия Фуатовна Столярова³

^{1,2,3} Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

¹ Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация

¹ aok@dscs.pro, <https://orcid.org/0000-0002-9811-5476>

² mva@dscs.pro✉, <https://orcid.org/0000-0002-5476-3025>

³ vfs@dscs.pro, <https://orcid.org/0000-0002-1666-2186>

Аннотация

Введение. Исследована возможность использования цифровых следов пользователей онлайн социальной сети, на примере тематики сообществ, для поддержки принятия решений в рамках профориентационной диагностики. **Метод.** Проведен статистический анализ: описательные статистики, z-критерий для сравнения двух групп и регрессионный анализ. В качестве цифровых следов пользователей проанализированы тематики подписок пользователей на различные сообщества, доступные в социальной сети, а также указанный в профиле пол респондента и количество друзей в социальной сети. Социально-профессиональная направленность личности оценена на основании результатов прохождения теста Голланда (под редакцией Г.В. Резапкиной). Выполнен поиск тенденций взаимосвязи между цифровыми следами пользователей, выраженными тематиками подписок, и ключевыми индикаторами социально-профессиональной направленности, отраженными в результатах теста Голланда, по данным пилотного исследования, проведенного с помощью приложения в онлайн социальной сети. **Основные результаты.** По результатам статистического анализа подтверждена гипотеза, что интересы пользователя, в виде тематик сообществ, взаимосвязаны с результатами теста Голланда. Доказаны гипотезы о существующих различиях в группах мужчин и женщин исследуемых признаков (результатов прохождения теста Голланда и ведущей тематики подписок на сообщество). При помощи регрессионного анализа среди группы женщин найдены взаимосвязи между преобладанием тематики сообществ «Образование» и ключевыми индикаторами: А (Художественный), Е (Предприимчивый), I (Интеллектуальный); преобладанием тематики «Стиль жизни» и выраженностью индикаторов: С (Конвенциональный), I, А, Е; «Средства массовой информации» и индикатором С. Среди группы мужчин была найдена взаимосвязь между преобладанием тематики «Спорт» и индикатором Е. **Обсуждение.** Результаты работы расширили пространство потенциальных предикторов профессиональной ориентации пользователей. Получена основа для проведения крупномасштабных исследований по квантификации и построению предиктивных моделей ключевых индикаторов профессий по тематикам подписок пользователей. Результаты могут быть использованы в направлении развития комплексного подхода к созданию рекомендательной системы для профессиональной ориентации пользователей.

Ключевые слова

профориентация, тест Голланда, профессиональные типы, выбор карьеры, социальные сети, анализ сообществ, интеллектуальная система, анализ цифровых следов

Благодарности

Работа выполнена в рамках проекта по государственному заданию Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» № FFZF-2022-0003 (проведение статистического анализа с применением критерия равенства пропорций для сравнения двух групп, регрессионный анализ); поддержана Санкт-Петербургским государственным университетом, проект № 75254082 (постановка задачи и гипотез исследования, обзор релевантных работ, описательные статистики).

© Хлобыстова А.О., Абрамов М.В., Столярова В.Ф., 2023

Ссылка для цитирования: Хлобыстова А.О., Абрамов М.В., Столярова В.Ф. Исследование тенденций взаимосвязи между профориентационными предпочтениями пользователей и их цифровыми следами в социальной сети // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 3. С. 564–574. doi: 10.17586/2226-1494-2023-23-3-564-574

Exploring the possibility of predicting users' career guidance preferences based on analysis of community topics and the gender in the online social network users' profiles

Anastasia O. Khlobystova¹, Maxim V. Abramov²✉, Valerie F. Stoliarova³

^{1,2,3} St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

¹ St. Petersburg State University (SPbSU), Saint Petersburg, 199034, Russian Federation

¹ aok@dscs.pro, <https://orcid.org/0000-0002-9811-5476>

² mva@dscs.pro✉, <https://orcid.org/0000-0002-5476-3025>

³ vfs@dscs.pro, <https://orcid.org/0000-0002-1666-2186>

Abstract

The possibility of using digital traces of online social network users, using community themes as an example, to support decision-making in career guidance diagnostics is investigated. Statistical analysis was performed: descriptive statistics, z-criterion for comparing two groups, and regression analysis. The themes of users' subscriptions to various communities available in the social network as well as the gender of the respondent and the number of friends in the social network indicated in the profile were analysed as digital user traces. The socio-professional orientation of the personality was assessed based on the results of the Holland test (edited by G.V. Rezapkina). The correlation between users' digital traces expressed by the themes of subscriptions, and key indicators of socio-professional orientation reflected in the results of the Holland test was analyzed based on the pilot study conducted through an online social networking application. The statistical analysis confirmed the hypothesis that user interests, in the form of community themes, are related to the results of the Holland test. The hypotheses of existing differences in the groups of men and women in the studied attributes (the results of the Holland test and the leading themes of community subscriptions) were proved. By means of regression analysis among the group of women the correlation was found between the prevalence of the community theme "Education" and the key indicators: A (Artistic), E (Enterprising), I (Intellectual); prevalence of the theme "Lifestyle" and severity of the indicators: C (Conventional), I, A, E; "Mass Media" and indicator C. Among the group of men, a correlation was found between the prevalence of "Sports" subject matter and indicator E. The results of the work expanded the space of potential predictors of users' vocational orientation. A foundation for large-scale research in quantifying and constructing predictive models of key occupational indicators based on users' subscription topics has been obtained. The results are useful in the direction of developing an integrated approach to creating a recommendation system for user career guidance.

Keywords

career guidance, Holland theory, RIASEC, vocational personality types, career path, online social networks, community analysis, intellectual system, digital footprint analysis

Acknowledgements

This work was carried out within the framework of the project under the state assignment of SPC RAS SPIIRAS No. FFZF-2022-0003 (performing statistical analysis using Z-criterion to compare two groups, regression analysis); with the financial support of St. Petersburg State University, project No. 75254082 (statement of the research problem and hypotheses, review of relevant papers, descriptive statistics).

For citation: Khlobystova A.O., Abramov M.V., Stoliarova V.F. Exploring the possibility of predicting users' career guidance preferences based on analysis of community topics and the gender in the online social network users' profiles. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 3, pp. 564–574 (in Russian). doi: 10.17586/2226-1494-2023-23-3-564-574

Введение

Согласно исследованию Росстата¹ до 57 % выпускников вузов устраиваются на работу по специальностям, не связанным с полученным образованием, а каждый четвертый россиянин не любит свою работу².

¹ Рабочая сила, занятость и безработица в России (по результатам выборочных обследований рабочей силы). 2022. Стат. сб./Росстат. М., 2022. 151 с. [Электронный ресурс]. Режим доступа: https://rosstat.gov.ru/storage/mediabank/Rab_sila_2022.pdf, свободный. Яз. рус. (дата обращения: 28.11.2022).

² Исследование показало, сколько россиян не любят свою работу [Электронный ресурс]. Режим доступа: <https://ria.ru/20220615/rabota-1795346061.html>, свободный. Яз. рус. (дата обращения: 28.11.2022).

При этом одной из ключевых карьерных ошибок является «неверный выбор профессии» (не соответствующий социально-профессиональной направленности личности)³. Таким образом, проблема профессионального самоопределения является острой на сегодняшний день.

³ Каждый восьмой россиянин считает своей главной карьерной ошибкой неправильный выбор профессии [Электронный ресурс]. Режим доступа: <https://spb.superjob.ru/research/articles/113577/kazhdyj-vosmoj-rossiyanin-schitaet-svoej-glavnoj-karernoj-oshibkoj-nepravilnyj-vybor-professii/>, свободный. Яз. рус. (дата обращения: 28.11.2022).

день. Традиционные подходы к решению обозначенной проблемы чаще всего основываются на консультациях с психологами или самостоятельном прохождении соискателем различных психологических тестов, ориентированных на его профессиональное самоопределение [1, 2]. Данные подходы наряду с преимуществами имеют недостатки: могут требовать существенных финансовых вложений; давать искаженный результат или быть неправильно интерпретированными. Отметим, что любое применение диагностических методик занимает большое время и может влиять на корректность конечного результата. Таким образом, актуален запрос на разработку эффективных подходов к профессиональной ориентации.

Проведен анализ современного состояния исследуемой области [3–28]. В результате не было обнаружено научных работ, в которых был бы осуществлен поиск связей между типом социально-профессиональной направленности и цифровыми следами пользователей, что подтвердило необходимость проведения настоящего исследования. Выявлена необходимость проверки гипотезы о том, что в качестве основного предиктора типа социально-профессиональной направленности личности могут выступать тематики подписок пользователей на различные сообщества, доступные в социальной сети «ВКонтакте». Выдвинутая гипотеза является новой и ранее в научных исследованиях не рассматривалась. Полномасштабное исследование, необходимое для построения предиктивной модели, в высокой степени ресурсозатратно, требует сбора большого набора данных, исследования зависимостей и построения модели. В связи с этим целесообразно провести пилотное исследование.

Цель работы — изучение тенденции взаимосвязи между цифровыми следами пользователей, выраженными тематиками подписок, и ключевыми индикаторами социально-профессиональной направленности, отраженными в результатах теста Голланда. Пилотное исследование выполнено с помощью приложения в онлайн социальной сети.

Научная новизна исследования заключается в расширении пространства потенциальных предикторов профессиональной ориентации пользователей. Теоретическая значимость — в развитии комплексного подхода, методов и моделей профессиональной ориентации пользователей через анализ их цифровых следов и последующего построения рекомендательной системы. Практическая значимость состоит в получении основы для проведения крупномасштабных исследований по квантификации и построению предиктивных моделей ключевых индикаторов профессий по тематикам подписок пользователей. Вместе с тем полученные результаты опосредованно вносят вклад в разработку практического инструментария, который направлен на снижение уровня безработицы и повышение уровня удовлетворенности жизнью посредством правильного выбора профессии.

Методы исследования

Одна из традиционных методик сопоставления психологического типа человека и его профессии

была предложена Дж. Голландом (модель RIASEC) [29]. Согласно теории Голланда, существует шесть типов личности и шесть профессиональных областей, с которыми данные типы личности соотносятся: R (Реалистичный), I (Интеллектуальный), A (Художественный), S (Социальный), E (Предприимчивый), C (Конвенциональный). Модель RIASEC адаптирована и локализована известным российским специалистом по вопросам профессионального самоопределения Г.В. Резапкиной [30]. В ходе прохождения теста Голланда респонденту необходимо было представить, что он обладает универсальными способностями и профессиональными навыками, и предложен список профессий, разбитых на пары. В каждой из пар необходимо выбрать профессию, которая кажется респонденту наиболее привлекательной. При этом каждая профессия сопоставлена одному из шести типов Голланда. Для получения результатов выполнен подсчет, сколько раз респондент выбрал профессию, относящуюся к тому или иному типу. Количество набранных баллов может быть разделено на следующие категории типов: 7–10 — ярко выраженный; 4–6 — средневыраженный; 0–3 — не выражен. Рассмотрим один ключевой индикатор социально-профессиональной направленности личности, т. е. тот тип, который набрал наибольшее число баллов.

Описание набора данных. Данные собраны на основании опроса, размещенного в веб-приложении на базе сервиса VK Mini Apps — «Психологические тесты»¹. На момент анализа результатов исследования тест Голланда прошли 2073 респондента. Тест мог пройти любой желающий, при этом программа собирала обезличенную информацию из аккаунта пользователей. В ходе предобработки данных были удалены незавершенные прохождения теста, а также закрытые аккаунты, из которых невозможно извлечь цифровые следы респондентов. В результате предобработки объем выборки сократился до 1237. Фрагмент полученного набора данных представлен в табл. 1, который состоит из: результатов прохождения теста Голланда: 6 типов личности A, C, E, I, R, S, находящихся в диапазоне от 0 до 13 (метрические переменные); информации о поле (1 — женский, 2 — мужской), указанном в профиле респондента (бинарная переменная); информацию о числе друзей в профиле (метрическая переменная); 28 переменных, соответствующих количеству категорий тематик сообществ «ВКонтакте», на которые подписан пользователь (метрические переменные).

Тест Голланда предполагает системную интерпретацию выраженности всех типов социально-профессиональной направленности личности. В рамках данной работы выбран один тип личности, который набрал наибольшее число баллов. Такой тип для каждого респондента назовем ключевым индикатором профессии, который определен на этапе предварительного

¹ Приложение «Психологические тесты» [Электронный ресурс]. Режим доступа: https://vk.com/services?w=app7794698_203437876, свободный. Яз. рус. (дата обращения: 25.11.2022).

Таблица 1. Фрагмент анализируемого набора данных
Table 1. Fragment of the analyzed data set

User_id	Типы личности						Пол sex	Количество друзей friends	Категории тематик сообществ «ВКонтакте»						
	A	C	E	I	R	S			auto	domestic services	cities & countries	...	hobby	connection serv	finance
1	5	6	6	4	10	11	2	137	1	0	0	...	3	0	1
2	9	7	11	3	7	5	1	22	3	0	9	...	8	0	12
3	8	6	6	7	9	6	1	24	0	0	9	...	1	0	0
4	7	11	4	5	7	8	2	287	0	2	8	...	25	0	7
5	7	2	10	8	7	8	2	404	1	0	5	...	11	0	1

анализа данных. Выполним анализ различия ключевых индикаторов профессии в гендерных группах и их взаимосвязь с ведущими тематиками среди подписок пользователей.

Основные результаты

Определение ведущих тематик сообществ. Сообщества в социальной сети «ВКонтакте» могут быть созданы по более чем 100 различным тематикам, которые разбиты на 28 укрупненных категорий. Например, категория «Дом, ремонт» включает такие тематики как «Дизайн интерьера», «Садоводство», «Строительство, ремонт». Для проведения анализа по каждому пользователю определим число подписок на сообщества по укрупненным категориям. В дальнейшем для простоты изложения текста отождествим понятия «тематика» и «укрупненная категория», подразумеваемая последнее. Для каждого пользователя определим *ведущую тематику*, которой посвящено наибольшее число подписок. Например, если из 28 рассматриваемых категорий у респондента 30 подписок на сообщества, относящиеся к категории «Красота и здоровье», 17 — «Медицина» и 5 — «Социальные организации», то ведущей для этого пользователя является категория «Красота и здоровье».

При проведении анализа данных выявлено, что для подавляющего числа респондентов (929 наблюдений) ведущими тематиками являются «Хобби» и «Развлечения». Отметим, что эти темы — зонтичные, включают в себя различные области повседневной жизни и во многом пересекаются с остальными тематиками. В итоге для пользователя выбирается ведущая тематика, имеющая наибольшее число подписок, кроме категорий «Хобби» и «Развлечения». Для описания сферы интереса респондентов сформируем две новые переменные: *isEnt* — бинарная переменная, принимающая истинное значение, если у пользователя ведущей тематикой являлось «Хобби» или «Развлечения»; *mainint* — категориальная переменная, определяющая основную тематику сообществ респондента, за исключением «Хобби» или «Развлечения».

Объединим малочисленные категории среди тематик, представляющих значения переменной *mainint*. Результат объединения представлен в табл. 2, объединенные тематики выделены жирным шрифтом.

Построим диаграмму размаха для исследуемых характеристик (рис. 1). При визуальном анализе диаграммы видно, что отдельные тематики ассоциированы с определенными профессиями. Численный анализ взаимосвязи ключевых индикаторов профессии и ведущих тематик осуществлен в рамках гендерных групп, указанных в профиле пользователя. Большая часть пользователей выборки отметили свой пол как женский (836 респондентов из 1237). Отметим, что соответствие указанного в профиле пола реальному полу пользователя онлайн социальной сети требует дополнительного анализа.

Далее проведем анализ, состоящий из двух этапов: исследуем статистические различия в ключевых индикаторах профессии и ведущих тематиках между группами мужчин и женщин; построим статистические регрессионные модели, позволяющие оценить качество взаимосвязи с учетом контекста, который отражает иные факторы, оказывающие влияние на результаты профориентационного теста. В настоящей работе контекстом выступают индикатор «Развлечения/Хобби», как первичной ведущей тематики, и число друзей пользователя в онлайн социальной сети.

Сравнение пропорций исследуемых признаков в группах мужчин и женщин. В рамках общей цели работы рассмотрим две статистические гипотезы:

- H1: ключевой тип социально-профессиональной направленности и пол независимы;
- H2: ведущая тематика сообществ и пол независимы.

Критерий проверки равенства пропорций используем для исследования гипотез H1 и H2. Сравнение пропорций, значения статистики критерия и соответствующее р-значение представлены в табл. 3 и 4.

Проведенный анализ не позволяет принять гипотезу H1 на уровне значимости 0,1, и, таким образом, можно говорить о взаимосвязи пола и ключевых индикаторов профессии. Среди мужчин, принявших участие в опросе, чаще в качестве ключевого типа профессионально-социальной направленности, встречаются E, R, при этом последний встречается наиболее часто; среди результатов женщин преобладают типы A и C.

Анализ ведущих предпочтений (тематик сообществ) не позволяет принять гипотезу H2 на уровне значимости 0,1. Можно сказать, что женщины более склонны к участию в сообществах с тематиками «Образование», «Домашнее хозяйство», «Образ жизни», а мужчины

Таблица 2. Сведения о ведущих тематиках респондентов
Table 2. Information on the leading themes of the respondents

Название переменной	Категории тематик, отнесенные к переменной	Число респондентов, для которых тематика ведущая
auto (спорт)	Авто, мото (auto) Спорт (sport)	78
lifestyle (образ жизни)	Красота, здоровье (beauty_health) Отношения, семья (relations) Медицина (medicine) Рестораны (restaurants)	261
cities (путешествия)	Города, страны (cities_countries) Туризм, путешествия и отдых (tourism)	77
computer_internet (цифровые технологии)	Компьютер, интернет	146
culture (искусство)	Культура, искусство (culture) Музыка (music)	70
education (образование)	Образование Финансы и страхование (finance)	168
household (домашнее хозяйство)	Животные (animals) Кулинария, рецепты (cooking_recipes) Дом, ремонт (house_repair) Профессиональные услуги (prof_service)	74
mass_media (средства массовой информации (СМИ))	СМИ (mass_media) Персоны (person) Социальные и общественные организации (social_orgs)	252
people_groups (группы по интересам)	Объединения, группы людей	52
shops (магазины)	Товары, магазины	43

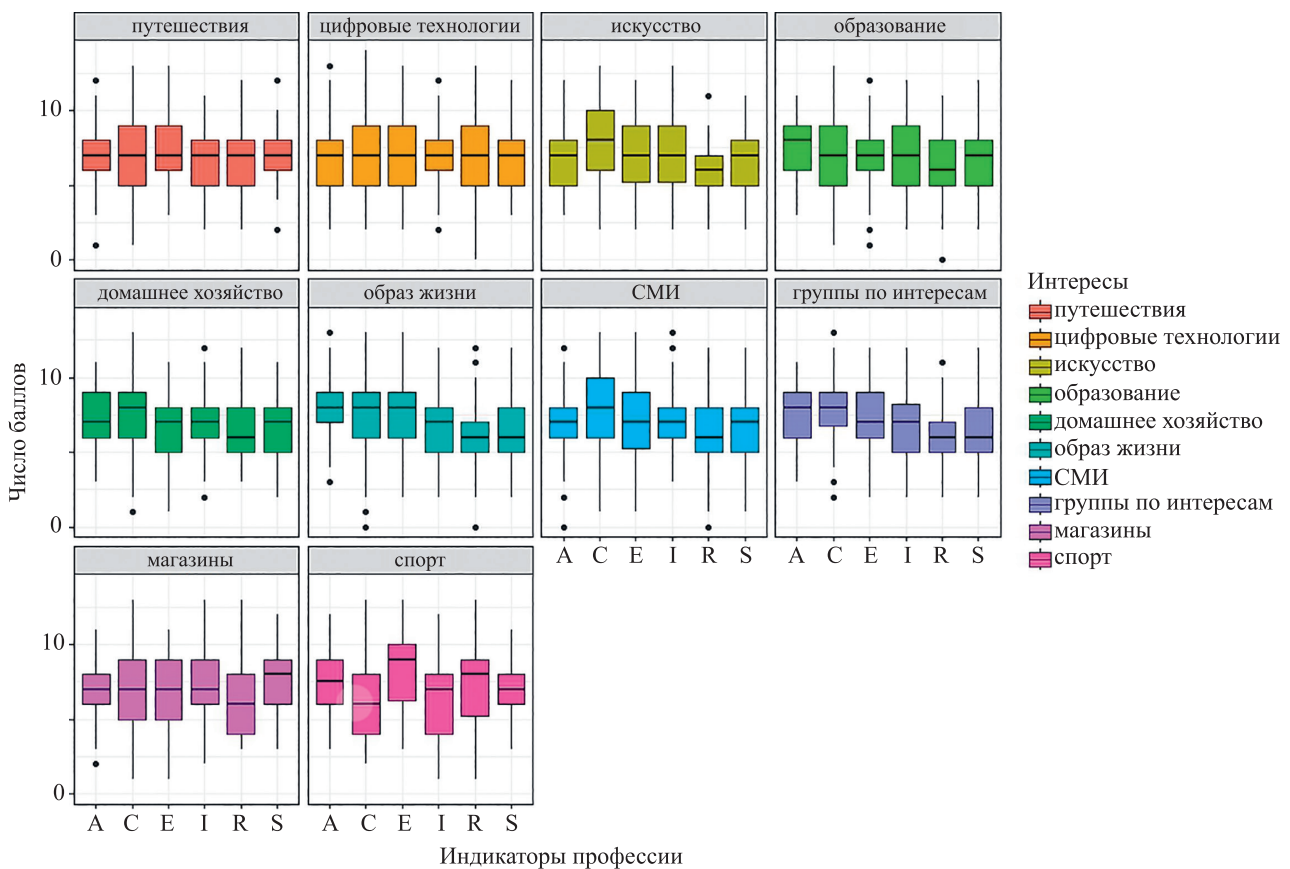


Рис. 1. Распределение индикаторов профессий по ведущим тематикам пользователей

Fig. 1. Distribution of occupational indicators by leading user topics

Таблица 3. Сравнение пропорций ключевых индикаторов в группах мужчин и женщин
Table 3. Comparison of the proportions of key indicators in groups of men and women

Наименование показателя	Ключевые индикаторы					
	A	C	E	I	R	S
Женский пол в профиле	0,17	0,29	0,17	0,2	0,09	0,08
Мужской пол в профиле	0,13	0,13	0,24	0,16	0,25	0,1
Значение статистики критерия	3,48 (1 df)	38,03 (1 df)	8,17 (1 df)	2,66 (1 df)	57,93 (1 df)	0,27 (1 df)
P-value	0,06	<0,05	<0,05	0,1	<0,05	0,61

Таблица 4. Сравнение ведущих тематик в группах мужчин и женщин
Table 4. Comparison of leading themes in groups of men and women

Ведущая тематика	Наименование показателя			
	Женский пол в профиле	Мужской пол в профиле	Значение chi-sq	P-value
Путешествия (cities)	0,07	0,05	0,51	0,47
Цифровые технологии (computer_internet)	0,06	0,25	84,26	<0,05
Культура (culture)	0,05	0,07	0,81	0,37
Образование (education)	0,16	0,09	12,23	<0,05
Домашнее хозяйство (household)	0,07	0,03	6,49	<0,05
Образ жизни (lifestyle)	0,3	0,03	107,16	<0,05
СМИ (mass_media)	0,18	0,25	7,41	<0,05
Группы по интересам (people_groups)	0,06	0,02	9,16	<0,05
Магазины (shops)	0,04	0,03	0,48	0,49
Спорт (sport)	0,01	0,18	121,92	<0,05

интересуются сообществами: «Цифровые технологии», «СМИ», «Спорт».

На следующем этапе анализа выполним построение регрессионных моделей. Модели позволят внутри каждой гендерной группы установить силу взаимосвязи между ключевыми индикаторами профессии и ведущими тематиками с учетом иных факторов, которые могут оказывать влияние на результаты проведения профориентационного тестирования.

Взаимосвязь ключевых индикаторов профессии и ведущих тематик по указанному в профиле пользователя социальной сети полу. В силу ограниченности числа наблюдений (836 женщин и 385 мужчин) и большого числа рассматриваемых ведущих тематик (10), в модель включим только следующие переменные: зависимые — ключевые индикаторы профессии (bestprof); независимые — переменную, отражающую ведущую тематику среди подписок респондента (education, lifestyle, mass_media, household, shops, computer_internet, cities, people_groups, culture, sport), переменную, показывающую является ли хобби или развлечения первостепенной ведущей тематикой (isEnt) и число друзей (friends).

Для квантификации взаимосвязи подгоним модели мультиномиальной логистической регрессии [31], которые представляют обобщение метода логистической регрессии на класс задач мультиномиальной классификации, где зависимая переменная представляет собой дискретную переменную с несколькими взаимоисключ

чающими категориями. В результате были подогнаны два набора моделей мультиномиальной регрессии, согласно указанному в профиле пользователя полу.

Значимость коэффициента при той или иной переменной определим согласно критерию отношения правдоподобия, который направлен на сравнение правдоподобий (likelihood) текущей модели с редуцированной, без рассматриваемой переменной.

Респонденты, у которых указан женский пол.

Для подгонки мультиномиальной регрессии используем пакет nnet [32] среды обработки данных R. На рис. 2 показаны графики зависимости вероятностей определенного индикатора профессии от рассматриваемых факторов (для каждой ведущей тематики отдельно). Полученные коэффициенты мультиномиальной логистической регрессии представлены по ссылке¹.

Отметим, что в нескольких категориях (culture, people_groups, shops, sport) выполнено небольшое число наблюдений, поэтому при построении регрессионных моделей они исключены из рассмотрения.

В результате анализа видно, что вероятность наличия ключевых индикаторов профессий C и I возрастает по сравнению с наличием индикатора A, при условии, что ведущая тематика респондента — «СМИ».

¹ Результаты мультиномиальной логистической регрессии [Электронный ресурс]. Режим доступа: <https://dscs.pro/materialy-po-proektam/rezultaty-issledovaniya-proforiyentatsiya/>, свободный. Яз. рус. (дата обращения: 17.01.2023).

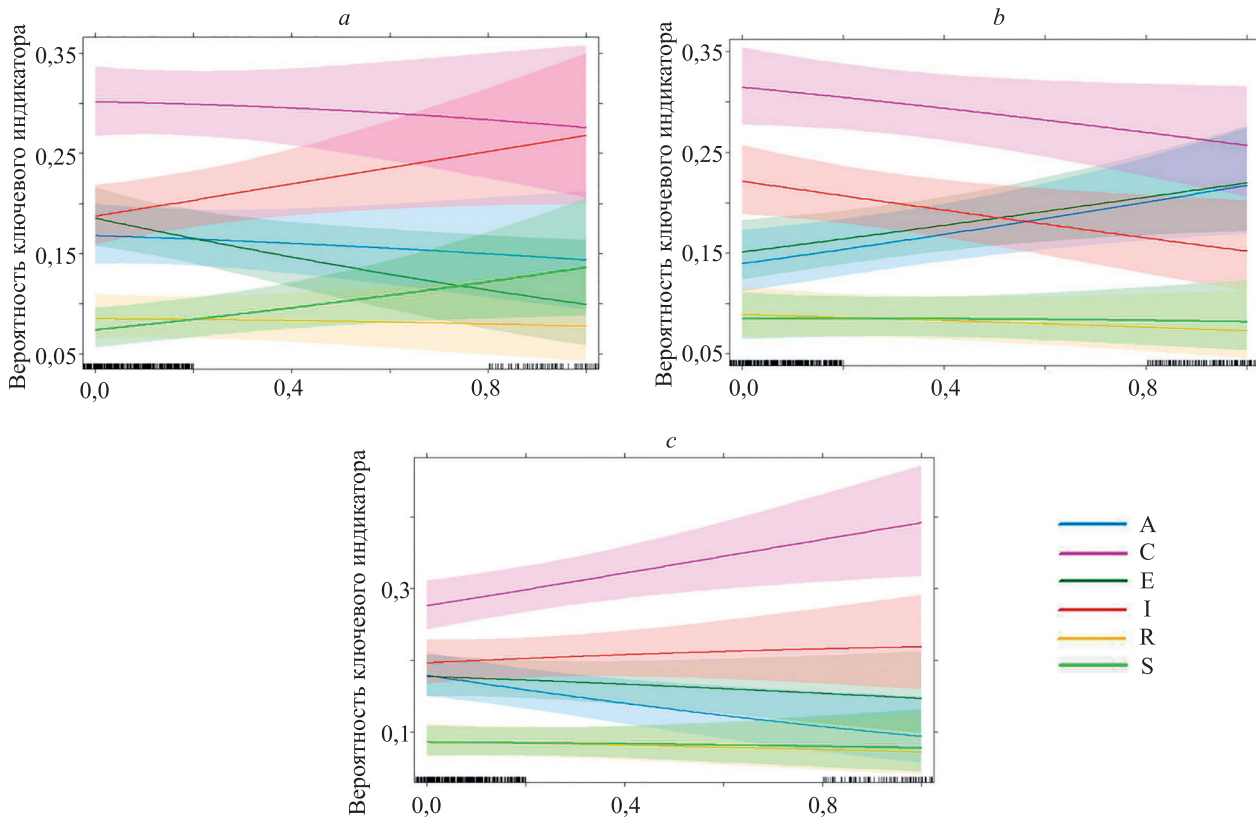


Рис. 2. Зависимости вероятностей наличия ключевых индикаторов профессии в результатах теста Голланда от наличия определенной ведущей тематики для респондентов, у которых указан женский пол в профиле: «Образование» (a); «Образ жизни» (b), «СМИ» (c)

Fig. 2. Dependencies of the probabilities of the presence of key occupational indicators in the results of the Holland test vs. the presence of a certain leading theme for respondents with a female gender in their profile: "Education" (a); "Lifestyle" (b), "Media" (c)

Если ведущая тематика — «Образ жизни», то вероятность наличия ключевых индикаторов C, I, R в результатах прохождения теста Голланда снижается, а если — «Образование», то вероятность индикатора S — повышается. Отметим, что увеличение числа друзей повышает вероятность в тесте ключевых индикаторов

профессии C, E, S по сравнению с вероятностью ключевого индикатора A.

Респонденты, у которых указан мужской пол.

Так как респондентов, у которых указан мужской пол, небольшое количество, то при построении регрессионных моделей используем только категории mass_media,

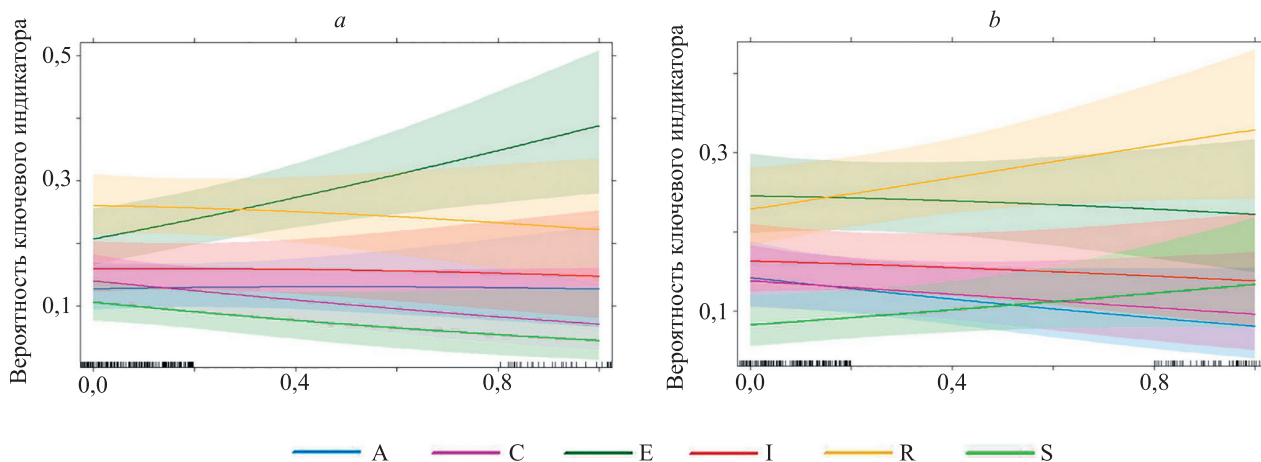


Рис. 3. Зависимости вероятностей наличия ключевых индикаторов профессии в результатах теста Голланда от ведущей тематики для респондентов, у которых указан мужской пол в профиле «Спорт» (a) и «Цифровые технологии» (b)

Fig. 3. Dependencies of the probabilities of the presence of key occupational indicators in Holland's test results vs. the leading theme for respondents who have a male gender in the profile "Sport" (a) and "Digital Technology" (b)

computer_internet и sport. Проведенный анализ показал, что наличие ведущей тематики «Спорт» увеличивает вероятность ключевого индикатора профессии E по сравнению с индикатором A в результатах теста Голланда у лиц, отметивших свой пол как мужской в профиле онлайн социальной сети (рис. 3, a). Наличие ведущей тематики «Цифровые технологии» увеличивает вероятность наличия ключевых индикаторов профессии R и S по сравнению с индикатором A (рис. 3, b), а наличие ведущей тематики «СМИ» увеличивает вероятность наличия ключевого индикатора C.

Выводы и обсуждение

По результатам статистического анализа была подтверждена гипотеза, что интересы пользователя, выражаемые в виде тематик сообществ, на которые он подписан в социальной сети «ВКонтакте», взаимосвязаны с результатами теста Голланда.

Отметим, что как ключевые индикаторы профессии, так и ведущие тематики имеют различия по полу (указанному в профиле онлайн социальной сети). Респонденты с указанным в аккаунте женским полом больше интересуются сообществами с тематиками «Образование», «Домашнее хозяйство», «Образ жизни», «Группы по интересам», с мужским полом — чаще имеют ведущую тематику «Цифровые технологии», «СМИ», «Спорт». Женщины более склонны иметь ключевые индикаторы профессии A и C, а мужчины — E и R. В силу небольшого числа наблюдений, а также, так как пол человека является важнейшим фактором, влияющим на проявление личностных особенностей, взаимосвязь ключевых индикаторов профессий и ведущих тематик рассмотрена отдельно по гендерным группам. Построение моделей мультимодальной регрессии показало следующие взаимосвязи. Для респондентов, у которых указан женский пол в профиле социальной сети, статистически значимую взаимосвязь удалось выявить между ведущей тематикой «Образование» и ключевым индикатором профессии S, тематикой «Образ жизни» и индикаторами C, I, R, тематикой «СМИ» и индикатором C. Кроме того, в этой группе статистически значимой является переменная, отражающая число друзей в профиле: увеличение числа друзей ведет к росту вероятности ключевых индикаторов профессии C, E, S в результатах теста Голланда.

Для группы мужчин не прослеживается подобной взаимосвязи ключевых индикаторов профессии с числом друзей. Кроме того, присутствует и иная картина ведущих тематик: в подавляющем большинстве случаев это или «СМИ», или «Спорт», или «Цифровые технологии». Была выявлена взаимосвязь тематики «Спорт» с ключевым индикатором профессии E, тематики «Цифровые технологии» с индикаторами R и S, тематики «СМИ» с ключевым индикатором C.

К ограничениям проведенного исследования относятся способ сбора данных, так как свободный доступ респондентов к приложению не позволяет оценить истинные характеристики выборки и определить подгруппу популяции, для которой получены результаты.

Кроме того, из-за малого количества наблюдений ряд ведущих тематик пришлось исключить из рассмотрения. Однако эта особенность может быть разрешена посредством рассмотрения иных способов моделирования, например, байесовского подхода к анализу данных за счет комбинирования информации из различных источников: статистической, экспертной и результатов предварительных исследований. Важной характеристикой таких моделей также является возможность обучения при поступлении новой информации, что позволит учитывать в финальных оценках статистической взаимосвязи исследуемых переменных новые результаты прохождения тестирований. Байесовская сеть доверия в данном случае будет представлена двумя взаимосвязанными блоками: блок аспектов социально-профессиональной направленности личности, отраженный в результатах теста Голланда (в том числе с учетом рангов всех шести измерений) и блок цифровых следов пользователя, включающий несколько ведущих тематик сообществ, пол, возраст, число друзей и другие. Таким образом, байесовские методы могут помочь смягчить проблему небольшого объема выборки, однако не отменяют необходимости проведения более масштабного исследования, направленного на построение предсказательной модели, осмысленность которого подтверждается в настоящей работе.

Также открытым вопросом исследования является и то, что профессиональные типы в чистом виде не встречаются [29, 30], т. е. можно говорить о ключевом типе социально-профессиональной направленности личности, что, в свою очередь, в дальнейшем позволит строить предположения о наиболее подходящей сфере деятельности. Вместе с тем для сужения области профессиональных интересов важно также рассмотреть выраженность второ- и третьестепенных индикаторов. Данный факт планируется учесть в дальнейших исследованиях, а именно, будет рассмотрен не один ключевой тип социально-профессиональной направленности, а совокупно нескольких преобладающих. Кроме того, для подтверждения результатов исследования в дальнейшем планируется расширить объем выборки и количество независимых переменных в модели мультиномиальной регрессии, в частности, такого рода увеличение числа переменных может основываться на восстановленной информации о профиле в социальных сетях: возрасте [33], интенсивности поведения [34], а также учитывать фактор влияния друзей и родственников пользователя [24, 25, 35].

Заключение

Изучены тенденции взаимосвязи между цифровыми следами пользователей, выраженными тематиками подписок, и ключевыми индикаторами социально-профессиональной направленности, отраженными в результатах теста Голланда, по данным пилотного исследования, проведенного с помощью приложения в онлайн социальной сети. При проведении исследования подтверждена гипотеза о том, что интересы пользователя, выраженные в виде тематик сообществ, на которые

он подписан в онлайн социальной сети, взаимосвязаны с результатами теста Голланда, т. е. могут выступать предикторами социально-профессиональной направленности личности. Полученные результаты позволили подтвердить целесообразность апробации примененных методик к более широкомасштабной выборке, а также учесть ряд тонких моментов при дизайне подобного рода исследования, в том числе сбор большего числа социально-демографических характеристик для определения подгруппы популяции и привлечение экспертов для укрупнения тематик сообществ. Результаты работы указали на наиболее перспективные направления исследований для определения взаимосвязи между интересами пользователей онлайн социальной сети и проявлениями типов социально-профессиональной направленности личности.

Проблема, рассмотренная в настоящей работе, отвечает одному из основных направлений научно-технологического развития Российской Федерации — «Созданию возможностей для выявления талантливой молодежи и построения успешной карьеры в области науки, технологий и инноваций, развитие интеллектуального потенциала страны»¹. Выполненное исследование дает основу для последующих разработок, ориентированных на снижение уровня безработицы и повышение уровня удовлетворенности жизнью посредством правильного своевременного выбора профессии.

¹ Кадры и человеческий капитал [Электронный ресурс]. Режим доступа: <https://xn--mlagf.xn--p1ai/ways/kadry-ichelovecheskiy-kapital/>, свободный. Яз. рус. (дата обращения: 28.11.2022).

Литература

- Xie D., Cobb C.L. Career and vocational interests and abilities, assessment of // *The Wiley Encyclopedia of Personality and Individual Differences: Models and Theories*. 2020. P. 493–497. <https://doi.org/10.1002/9781118970843.ch152>
- Malik-Liévano B., Álvarez-González B., Sánchez-García M.F., Irving B.A. *International Perspectives on Research in Educational and Career Guidance*. Springer Cham, 2020. 173 p. <https://doi.org/10.1007/978-3-030-26135-1>
- Bogacheva E., Tatarenko F., Smetannikov I. Predicting vocational personality type from socio-demographic features using machine learning methods // *CCRS '20: Proc. of the 2020 1st International Conference on Control, Robotics and Intelligent System*. 2020. P. 93–98. <https://doi.org/10.1145/3437802.3437819>
- Сербина Г.Н., Мацута В.В., Гойко В.Л. Анализ связи психологических характеристик пользователей социальной сети "ВКонтакте" с подписками на сообщества с девиантным контентом // *Вестник Томского государственного университета*. 2021. № 467. С. 164–169. <https://doi.org/10.17223/15617793/467/20>
- Мацута В.В., Гойко В.Л., Фещенко А.В., Уляков М.А., Петров Е.Ю. Платформа психологической диагностики с авторизацией респондентов в социальной сети "ВКонтакте". Свидетельство государственной регистрации программы для ЭВМ № 2021612933. Бюл. 2021. № 3.
- Dağıstanlı Ö., Erbay H., Kör H., Yurtakal A.H. Reflection of people's professions on social media platforms // *Neural Computing and Applications*. 2023. V. 35. N 7. P. 5575–5586. <https://doi.org/10.1007/s00521-022-07987-8>
- Van Huizen P., Mason R., Williams B. Exploring paramedicine student preferences using Holland's vocational theory: a cross-sectional study // *Nursing & Health Sciences*. 2021. V. 23. N 4. P. 818–824. <https://doi.org/10.1111/nhs.12870>
- Ding Y., Wang Q., Hourieh N., Yu Q. Vocational personality types in college engineering students in relation to academic achievement // *Journal of Employment Counseling*. 2020. V. 57. N 1. P. 27–47. <https://doi.org/10.1002/joec.12136>
- Earl J., Iskandar F., Elizondo F. Take a job, any job: Exploring the importance of matched interests to career paths and work satisfaction // *Journal of Employment Counseling*. 2019. V. 56. N 1. P. 33–45. <https://doi.org/10.1002/joec.12101>
- Gunawardena A.D.A., Meyer R.R., Kularatna T., Samaranyake S., Gunawardena R.H. Interest-aligned system for college degree planning and career path selection // *SN Computer Science*. 2022. V. 3. N 3. P. 221. <https://doi.org/10.1007/s42979-022-01107-y>
- Su J.-H., Liao Y.-W., Xu J.-Z., Zhao Y.-W. A personality-driven recommender system for cross-domain learning based on Holland code assessments // *Sustainability*. 2021. V. 13. N 7. P. 3936. <https://doi.org/10.3390/su13073936>
- Shminan A.S., Choi L.J., Barawi M.H., Hashim W.N.W., Andy H. InVesa 1.0: The conceptual framework of interactive virtual academic advisor system based on psychological profiles // *Proc. of the 13th International Conference on Information & Communication*

References

- Xie D., Cobb C.L. Career and vocational interests and abilities, assessment of. *The Wiley Encyclopedia of Personality and Individual Differences: Models and Theories*, 2020, pp. 493–497. <https://doi.org/10.1002/9781118970843.ch152>
- Malik-Liévano B., Álvarez-González B., Sánchez-García M.F., Irving B.A. *International Perspectives on Research in Educational and Career Guidance*. Springer Cham, 2020, 173 p. <https://doi.org/10.1007/978-3-030-26135-1>
- Bogacheva E., Tatarenko F., Smetannikov I. Predicting vocational personality type from socio-demographic features using machine learning methods. *CCRS '20: Proc. of the 2020 1st International Conference on Control, Robotics and Intelligent System*, 2020, pp. 93–98. <https://doi.org/10.1145/3437802.3437819>
- Serbina G.N., Matsuta V.V., Goiko V.L. Analysis of the relationship between VK users' psychological characteristics and subscriptions to communities with deviant content. *Tomsk State University Journal*, 2021, no. 467, pp. 164–169. (in Russian). <https://doi.org/10.17223/15617793/467/20>
- Matcuta V.V., Goiko V.L., Feshchenko A.V., Uliakov M.A., Petrov E. Yu. Platform for psychological diagnostics with authorization of respondents in the social network "VKontakte". *Certificate of the state computer program registration 2021612933*. 2021. (in Russian)
- Dağıstanlı Ö., Erbay H., Kör H., Yurtakal A.H. Reflection of people's professions on social media platforms. *Neural Computing and Applications*, 2023, vol. 35, no. 7, pp. 5575–5586. <https://doi.org/10.1007/s00521-022-07987-8>
- Van Huizen P., Mason R., Williams B. Exploring paramedicine student preferences using Holland's vocational theory: a cross-sectional study. *Nursing & Health Sciences*, 2021, vol. 23, no. 4, pp. 818–824. <https://doi.org/10.1111/nhs.12870>
- Ding Y., Wang Q., Hourieh N., Yu Q. Vocational personality types in college engineering students in relation to academic achievement. *Journal of Employment Counseling*, 2020, vol. 57, no. 1, pp. 27–47. <https://doi.org/10.1002/joec.12136>
- Earl J., Iskandar F., Elizondo F. Take a job, any job: Exploring the importance of matched interests to career paths and work satisfaction. *Journal of Employment Counseling*, 2019, vol. 56, no. 1, pp. 33–45. <https://doi.org/10.1002/joec.12101>
- Gunawardena A.D.A., Meyer R.R., Kularatna T., Samaranyake S., Gunawardena R.H. Interest-aligned system for college degree planning and career path selection. *SN Computer Science*, 2022, vol. 3, no. 3, pp. 221. <https://doi.org/10.1007/s42979-022-01107-y>
- Su J.-H., Liao Y.-W., Xu J.-Z., Zhao Y.-W. A personality-driven recommender system for cross-domain learning based on Holland code assessments. *Sustainability*, 2021, vol. 13, no. 7, pp. 3936. <https://doi.org/10.3390/su13073936>
- Shminan A.S., Choi L.J., Barawi M.H., Hashim W.N.W., Andy H. InVesa 1.0: The conceptual framework of interactive virtual academic advisor system based on psychological profiles. *Proc. of the 13th International Conference on Information & Communication*

- Technology and System (ICTS). 2021. P. 112–117. <https://doi.org/10.1109/ICTS52701.2021.9608182>
13. Usslepp N., Hübner N., Stoll G., Spengler M., Trautwein U., Nagengast B. RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices? // *Journal of Personality*. 2020. V. 88. N 5. P. 1007–1024. <https://doi.org/10.1111/jopy.12547>
 14. Cioca L.-I., Bratu M.L. Sustainability of youth careers in Romania—Study on the correlation of students' personal interests with the selected university field of study // *Sustainability*. 2020. V. 13. N 1. P. 229. <https://doi.org/10.3390/su13010229>
 15. Ethan Daphne C., Stanley M., Augustine V., Anggreainy M.S., Setiawan A.S. Developing an application to recommend major based on personality test using waterfall method // *Proc. of the 5th International Conference of Computer and Informatics Engineering (IC2IE)*. 2022. P. 220–224. <https://doi.org/10.1109/IC2IE56416.2022.9970189>
 16. Rocconi L.M., Liu X., Pike G.R. The impact of person-environment fit on grades, perceived gains, and satisfaction: an application of Holland's theory // *Higher Education*. 2020. V. 80. N 5. P. 857–874. <https://doi.org/10.1007/s10734-020-00519-0>
 17. Sharma S., Behl R. Analysing the impact of social media on students' academic performance: A comparative study of extraversion and introversion personality // *Psychological Studies*. 2022. V. 67. N 4. P. 549–559. <https://doi.org/10.1007/s12646-022-00675-6>
 18. Gonçalves A.P., Martins G.H., Salvador A.P., Machado G.M., Carvalho L.D.F., Ambiel R.A.M. Career adaptability and associations with personality traits and pathological traits utilizing network analysis // *International Journal for Educational and Vocational Guidance*. 2021. V. 21. N 2. P. 379–394. <https://doi.org/10.1007/s10775-020-09448-x>
 19. Kircaburun K., Alhabash S., Tosuntaş Ş.B., Griffiths M.D. Uses and gratifications of problematic social media use among university students: A simultaneous examination of the Big Five of personality traits, social media platforms, and social media use motives // *International Journal of Mental Health and Addiction*. 2020. V. 18. N 3. P. 525–547. <https://doi.org/10.1007/s11469-018-9940-6>
 20. Dandannavar P.S., Mangalwede S.R., Kulkarni P.M. Predicting the primary dominant personality trait of perceived leaders by mapping linguistic cues from social media data onto the big five model // *Advances in Intelligent Systems and Computing*. 2021. V. 1141. P. 417–428. https://doi.org/10.1007/978-981-15-3383-9_37
 21. Biswas K., Shivakumara P., Pal U., Lu T. A new ontology-based multimodal classification system for social media images of personality traits // *Signal, Image and Video Processing*. 2023. V. 17. N 2. P. 543–551. <https://doi.org/10.1007/s11760-022-02259-3>
 22. Hoff K.A., Chu C., Einarsdóttir S., Briley D.A., Hanna A., Rounds J. Adolescent vocational interests predict early career success: Two 12-year longitudinal studies // *Applied Psychology*. 2022. V. 71. N 1. P. 49–75. <https://doi.org/10.1111/apps.12311>
 23. Oliveira Í.M., Porfeli E.J., do Céu Taveira M., Lee B. Children's career expectations and parents' jobs: intergenerational (dis)continuities // *The Career Development Quarterly*. 2020. V. 68. N 1. P. 63–77. <https://doi.org/10.1002/cdq.12213>
 24. Banov K., Krapic N., Kardum I. Do vocational interests matter for the selection of romantic partners? Evidence from variable- and couple-centered approaches // *Applied Psychology*. 2023. V. 72. N 2. P. 697–717. <https://doi.org/10.1111/apps.12396>
 25. Mayrand K., Sabourin S., Savard C. Vocational personality in romantic relationships: how vocational similarity contributes to couple functioning // *Contemporary Family Therapy*. 2023. V. 45. N 2. P. 172–185. <https://doi.org/10.1007/s10591-021-09620-7>
 26. Kim J., Park D., Shin Y.J. Friendship dynamics of career decision-making self-efficacy: a longitudinal social network approach // *Current Psychology*. 2022. in press. <https://doi.org/10.1007/s12144-022-03887-1>
 27. Song Q.C., Shin H.J., Tang C., Hanna A., Behrend T. Investigating machine learning's capacity to enhance the prediction of career choices // *Personnel Psychology*. 2022. in press. <https://doi.org/10.1111/peps.12529>
 28. José-García A., Sneyd A., Melro A., Ollagnier A., Tarling G., Zhang H., Stevenson M., Everson R., Arthur R. C3-IoC: A career guidance system for assessing student skills using machine learning and network visualisation // *International Journal of Artificial Intelligence in Education*. 2022. in press. <https://doi.org/10.1007/s40593-022-00317-y>
- Technology and System (ICTS)*, 2021, pp. 112–117. <https://doi.org/10.1109/ICTS52701.2021.9608182>
13. Usslepp N., Hübner N., Stoll G., Spengler M., Trautwein U., Nagengast B. RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices? // *Journal of Personality*, 2020, vol. 88, no. 5, pp. 1007–1024. <https://doi.org/10.1111/jopy.12547>
 14. Cioca L.-I., Bratu M.L. Sustainability of youth careers in Romania—Study on the correlation of students' personal interests with the selected university field of study. *Sustainability*, 2020, vol. 13, no. 1, pp. 229. <https://doi.org/10.3390/su13010229>
 15. Ethan Daphne C., Stanley M., Augustine V., Anggreainy M.S., Setiawan A.S. Developing an application to recommend major based on personality test using waterfall method. *Proc. of the 5th International Conference of Computer and Informatics Engineering (IC2IE)*, 2022, pp. 220–224. <https://doi.org/10.1109/IC2IE56416.2022.9970189>
 16. Rocconi L.M., Liu X., Pike G.R. The impact of person-environment fit on grades, perceived gains, and satisfaction: an application of Holland's theory. *Higher Education*, 2020, vol. 80, no. 5, pp. 857–874. <https://doi.org/10.1007/s10734-020-00519-0>
 17. Sharma S., Behl R. Analysing the impact of social media on students' academic performance: A comparative study of extraversion and introversion personality. *Psychological Studies*, 2022, vol. 67, no. 4, pp. 549–559. <https://doi.org/10.1007/s12646-022-00675-6>
 18. Gonçalves A.P., Martins G.H., Salvador A.P., Machado G.M., Carvalho L.D.F., Ambiel R.A.M. Career adaptability and associations with personality traits and pathological traits utilizing network analysis. *International Journal for Educational and Vocational Guidance*, 2021, vol. 21, no. 2, pp. 379–394. <https://doi.org/10.1007/s10775-020-09448-x>
 19. Kircaburun K., Alhabash S., Tosuntaş Ş.B., Griffiths M.D. Uses and gratifications of problematic social media use among university students: A simultaneous examination of the Big Five of personality traits, social media platforms, and social media use motives. *International Journal of Mental Health and Addiction*, 2020, vol. 18, no. 3, pp. 525–547. <https://doi.org/10.1007/s11469-018-9940-6>
 20. Dandannavar P.S., Mangalwede S.R., Kulkarni P.M. Predicting the primary dominant personality trait of perceived leaders by mapping linguistic cues from social media data onto the big five model. *Advances in Intelligent Systems and Computing*, 2021, vol. 1141, pp. 417–428. https://doi.org/10.1007/978-981-15-3383-9_37
 21. Biswas K., Shivakumara P., Pal U., Lu T. A new ontology-based multimodal classification system for social media images of personality traits. *Signal, Image and Video Processing*, 2023, vol. 17, no. 2, pp. 543–551. <https://doi.org/10.1007/s11760-022-02259-3>
 22. Hoff K.A., Chu C., Einarsdóttir S., Briley D.A., Hanna A., Rounds J. Adolescent vocational interests predict early career success: Two 12-year longitudinal studies. *Applied Psychology*, 2022, vol. 71, no. 1, pp. 49–75. <https://doi.org/10.1111/apps.12311>
 23. Oliveira Í.M., Porfeli E.J., do Céu Taveira M., Lee B. Children's career expectations and parents' jobs: intergenerational (dis)continuities. *The Career Development Quarterly*, 2020, vol. 68, no. 1, pp. 63–77. <https://doi.org/10.1002/cdq.12213>
 24. Banov K., Krapic N., Kardum I. Do vocational interests matter for the selection of romantic partners? Evidence from variable- and couple-centered approaches. *Applied Psychology*, 2023, vol. 72, no. 2, pp. 697–717. <https://doi.org/10.1111/apps.12396>
 25. Mayrand K., Sabourin S., Savard C. Vocational personality in romantic relationships: how vocational similarity contributes to couple functioning. *Contemporary Family Therapy*, 2023, vol. 45, no. 2, pp. 172–185. <https://doi.org/10.1007/s10591-021-09620-7>
 26. Kim J., Park D., Shin Y.J. Friendship dynamics of career decision-making self-efficacy: a longitudinal social network approach. *Current Psychology*, 2022, in press. <https://doi.org/10.1007/s12144-022-03887-1>
 27. Song Q.C., Shin H.J., Tang C., Hanna A., Behrend T. Investigating machine learning's capacity to enhance the prediction of career choices. *Personnel Psychology*, 2022, in press. <https://doi.org/10.1111/peps.12529>
 28. José-García A., Sneyd A., Melro A., Ollagnier A., Tarling G., Zhang H., Stevenson M., Everson R., Arthur R. C3-IoC: A career guidance system for assessing student skills using machine learning and network visualisation. *International Journal of Artificial Intelligence in Education*, 2022, in press. <https://doi.org/10.1007/s40593-022-00317-y>

29. Holland J.L. Explorations of a theory of vocational choice and achievement: II. A four-year prediction study // *Psychological Reports*. 1963. V. 12. N 2. P. 547–594. <https://doi.org/10.2466/pr0.1963.12.2.547>
30. Резапкина Г.В. Психология и выбор профессии: программа предпрофильной подготовки. Учебно-методическое пособие для психологов и педагогов. М.: Генезис, 2005. 208 с.
31. Agresti A. *Categorical Data Analysis*. John Wiley & Sons, 2003.
32. Venables W.N., Ripley B.D. *Modern Applied Statistics with S*. Fourth Edition. New York: Springer, 2002. 498 p.
33. Корепанова А.А., Абрамов М.В. Применение случайного леса в выборе метода восстановления возраста пользователя социальной сети // *Искусственный интеллект и принятие решений*. 2021. № 2. С. 66–76. <https://doi.org/10.14357/20718594210207>
34. Торопова А.В., Абрамов М.В., Тулупьева Т.В. Машинное обучение байесовской сети доверия как инструмента оценки интенсивности процесса по данным из социальной сети // *Научно-технический вестник информационных технологий, механики и оптики*. 2021. Т. 21. № 5. С. 727–737. <https://doi.org/10.17586/2226-1494-2021-21-5-727-737>
35. Khlobystova A., Abramov M., Korepanova A., Liapin N. Identification of predictors for estimation the intensity of relationships between users of online social networks // *Lecture Notes in Networks and Systems*. 2023. V. 566. P. 216–225. https://doi.org/10.1007/978-3-031-19620-1_21
29. Holland J.L. Explorations of a theory of vocational choice and achievement: II. A four-year prediction study. *Psychological Reports*, 1963, vol. 12, no. 2, pp. 547–594. <https://doi.org/10.2466/pr0.1963.12.2.547>
30. Rezapkina G.V. *Psychology and Profession Selection: Pre-Profile Training Program. Study Guide for psychologists and teachers*. Moscow, Genesis Publ., 2005, 208 p. (in Russian)
31. Agresti A. *Categorical Data Analysis*. John Wiley & Sons, 2003.
32. Venables W.N., Ripley B.D. *Modern Applied Statistics with S*. Fourth Edition. New York, Springer, 2002, 498 p.
33. Korepanova A.A., Abramov M.V. Application of random forest in choosing a method of recovering the age of social network users. *Scientific and Technical Information Processing*, 2022, vol. 49, no. 5, pp. 317–324. <https://doi.org/10.3103/s0147688222050057>
34. Toropova A.V., Abramov M.V., Tulupyeva T.V. Machine learning of the bayesian belief network as a tool for evaluating the process frequency on social network data. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 5, pp. 727–737. (in Russian). <https://doi.org/10.17586/2226-1494-2021-21-5-727-737>
35. Khlobystova A., Abramov M., Korepanova A., Liapin N. Identification of predictors for estimation the intensity of relationships between users of online social networks. *Lecture Notes in Networks and Systems*, 2023, vol. 566, pp. 216–225. https://doi.org/10.1007/978-3-031-19620-1_21

Авторы

Хлобystова Анастасия Олеговна — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация; ассистент, Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация, [sc 57204970981](https://orcid.org/0000-0002-9811-5476), <https://orcid.org/0000-0002-9811-5476>, aok@dscs.pro

Абрамов Максим Викторович — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 56938320500](https://orcid.org/0000-0002-5476-3025), <https://orcid.org/0000-0002-5476-3025>, mva@dscs.pro

Столярова Валерия Фуатовна — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57220208479](https://orcid.org/0000-0002-1666-2186), <https://orcid.org/0000-0002-1666-2186>, vfs@dscs.pro

Статья поступила в редакцию 28.12.2022
Одобрена после рецензирования 27.02.2023
Принята к печати 26.05.2023

Authors

Anastasia O. Khlobystova — Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation; Assistant, St. Petersburg State University (SPbSU), Saint Petersburg, 199034, Russian Federation, [sc 57204970981](https://orcid.org/0000-0002-9811-5476), <https://orcid.org/0000-0002-9811-5476>, aok@dscs.pro

Maxim V. Abramov — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 56938320500](https://orcid.org/0000-0002-5476-3025), <https://orcid.org/0000-0002-5476-3025>, mva@dscs.pro

Valerie F. Stolarova — Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 57220208479](https://orcid.org/0000-0002-1666-2186), <https://orcid.org/0000-0002-1666-2186>, vfs@dscs.pro

Received 28.12.2022
Approved after reviewing 27.02.2023
Accepted 26.05.2023



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»