

Онлайн-образование после пандемии: исследование проблем и возможностей студентов инструментами больших данных

Научная статья

DOI: 10.31992/0869-3617-2023-32-10-133-150

Богданова Анна Владимировна – канд. пед. наук, начальник отдела технологий онлайн-образования, ORCID: 0000-0002-3553-2272, a.bogdanova@tltso.ru

Тольяттинский государственный университет, г. Тольятти, Российская Федерация

Адрес: 445020, Самарская обл., г. Тольятти, Белорусская ул., 14;

Александрова Юлия Константиновна – мл. науч. сотрудник Центра прикладного анализа больших данных, ORCID: 0000-0002-6069-779X, jalexandrova@data.tsu.ru

Томский государственный университет, г. Томск, Российская Федерация

Адрес: 634050, Томск, пр-т Ленина, 36;

Гойко Вячеслав Леонидович – заведующий научно-исследовательской лаборатории прикладного анализа больших данных, ORCID: 0000-0002-5985-3724, goiko@data.tsu.ru

Томский государственный университет, г. Томск, Российская Федерация

Адрес: 634050, Томск, пр-т Ленина, 36;

Орлова Вера Вениаминовна – д-р социол. наук, проф., заведующая кафедрой философии и социологии, ORCID: 0000-0001-6617-5346, SPIN-код: 6305-8140, orlova_vv@mail.ru

Томский государственный университет систем управления и радиоэлектроники, г. Томск, Российская Федерация

Адрес: 634034, Томск, пр-т Ленина, 40.

Аннотация. В статье представлен научно обоснованный подход к анализу больших объёмов данных цифровых следов студентов в социальных сетях, который позволяет эффективно выявлять возникающие и наиболее обсуждаемые проблемы у студентов, а также выделять болевые точки, дающие возможности роста, развития вузов и улучшения характеристик образовательного процесса, сопровождения студентов и т. д. Исследование основано на тематическом анализе сообщений, опубликованных в университетских сообществах в социальной сети «ВКонтакте» инструментами больших данных. Результаты исследования показали, что студенты российских вузов до сих пор сталкиваются с рядом проблем, включая слабую техническую инфраструктуру университетов, цифровое неравенство в доступе к онлайн-образованию и негативное отношение к дистанционному обучению.

Научная проблема исследования заключается в противоречии между существующим объёмом неструктурированных данных цифровых следов студентов в социальных сетях и отсутствием научно-обоснованного и апробированного методологического подхода к анализу и оценке этих объёмных данных, что создаёт препятствия для фундаментального исследования взаимосвязи между активностью студентов в социальных сетях и их удовлетворён-

ностью качеством образовательного процесса. Практическая направленность определяется в проведении анализа данных с применением инструментов больших данных. Полученные результаты и научно обоснованные выводы полезны для разработки инновационных стратегий и инструментов оценки и поддержки студентов.

Результаты показывают, что отслеживание трендов на основе цифровых следов студентов в социальных сетях и инструментария больших данных даёт высокую точность аналитических данных и может стать основой для выявления проблемных ситуаций в отдельных вузах и отрасли в целом, для принятия решений и управления, основанного на данных.

Ключевые слова: удовлетворённость студентов, высшее образование, интеллектуальный анализ данных, большие данные, онлайн-образование, качество образования, цифровой след, социальные сети

Для цитирования: Богданова А.В., Александрова Ю.К., Гойко В. А., Орлова В.В. Онлайн-образование после пандемии: исследование проблем и возможностей студентов инструментами больших данных // Высшее образование в России. 2023. Т. 32. № 10. С. 133–150. DOI: 10.31992/0869-3617-2023-32-10-133-150

Online Education after the Pandemic: Student Problems and Opportunities Research Using Big Data Tools

Original article

DOI: 10.31992/0869-3617-2023-32-10-133-150

Anna V. Bogdanova – Cand. Sci. (Pedagogical Sciences), Head of the Department of Online Education Technologies, ORCID: 0000-0002-3553-2272, a.bogdanova@tltsu.ru

Togliatti State University, Togliatti, Russian Federation

Address: 14 Belorusskaya str., Togliatti, 445020, Samara region, Russian Federation;

Yuliya K. Aleksandrova – Junior Research Fellow, Center for Applied Big Data Analysis, ORCID: 0000-0002-6069-779X, jalexandrova@data.tsu.ru

Tomsk State University, Tomsk, Russian Federation

Address: 36 Lenina ave., Tomsk, 634050, Russian Federation;

Vyacheslav L. Goiko – Head of the Research Laboratory for Applied Big Data Analysis, ORCID: 0000-0002-5985-3724, goiko@data.tsu.ru

Tomsk State University, Tomsk, Russian Federation

Address: 36 Lenina ave., Tomsk, 634050, Russian Federation;

Vera V. Orlova – Dr. Sci. (Sociology), Professor of the Department of Philosophy and Sociology, ORCID 0000-0001-6617-5346, SPIN-код: 6305-8140, orlova_vv@mail.ru

Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

Address: 40 Lenina ave., Tomsk, 634034, Russian Federation

Abstract. This paper presents a scientifically based approach to analyzing large volumes of data from digital traces of students on social networks, which allows you to effectively identify emerging and most discussed problems among students, as well as highlight pain points that provide opportunities for growth, development of universities and improvement of the characteristics of the educational process, support for students etc. The study is based on a thematic analysis of messages

published in university communities on the VKontakte social network using big data tools. The study results showed that Russian university students still face a number of challenges, including weak technical infrastructure at universities, a digital divide in access to online education, and negative attitudes towards distance learning.

The scientific problem of the study is the contradiction between the existing volume of unstructured data of students' digital traces in social networks and the lack of a scientifically-based and proven methodological approach to the analysis and evaluation of this voluminous data, which creates obstacles to fundamental research into the relationship between students' activity in social networks and their satisfaction quality of the educational process. The practical focus is determined in conducting data analysis using big data tools. The findings and evidence-based implications are useful for developing innovative strategies and tools for assessing and supporting students.

The results show that the use of big data tools for tracking trends based on digital traces of students on social networks provides highly accurate analytical data and can become the basis for identifying problematic situations in individual universities and the industry as a whole, for data-driven decision-making and management.

Keywords: student satisfaction, higher education, data mining, big data, online education, education quality, digital footprint, social networks.

Cite as: Bogdanova, A.V., Aleksandrova, Yu.K., Goiko, V.L., Orlova, V.V. (2023). Online Education after the Pandemic: Student Problems and Opportunities Research Using Big Data Tools. *Vysshee obrazovanie v Rossii = Higher Education in Russia*. Vol. 32, no. 10, pp. 133-150, doi: 10.31992/0869-3617-2023-32-10-133-150 (In Russ., abstract in Eng.).

Введение

Высшее образование оказывает существенное влияние на экономическое развитие большинства стран мира [1; 2]. Образовательная успешность студентов, которая в значительной степени зависит от условий обучения, при этом является важным фактором, оказывающим влияние на экономическую и социальную эффективность высшего образования [3; 4]. Именно поэтому ориентация на студента (студентоцентричность), учёт его потребностей и удовлетворённости учебным процессом сегодня всё чаще становятся предметом внимания исследователей [5].

Прошло уже более трёх лет с тех пор, как 11 марта 2020 г. Всемирная организация здравоохранения (ВОЗ) объявила вспышку COVID-19 пандемией. На этом фоне развернулся глобальный образовательный кризис: в начале апреля 2020 г. в 194 странах были закрыты все учебные заведения, что затронуло

почти 1,6 миллиарда учащихся и студентов, что составило 91,3% от их общемирового числа [6]. Никогда прежде не происходил такой быстрый и масштабный переход образования в онлайн-формат, что коренным образом изменило характер и условия преподавания и обучения.

Сильными сторонами в ситуации экстренного массового перехода к дистанционному обучению в России стали:

- доступный Интернет: в 2022 г. Россия заняла 6-е место рейтинга стран с самым доступным Интернетом¹, что даёт возможность удалённого доступа к электронным контентам и онлайн-мероприятиям широкому кругу граждан;
- развитая сеть партнёрств между академической и профессиональной сферами, сотрудничество между университетами и отраслевыми предприятиями, возможность удалённой организации практики и стажировки.

¹ Россия заняла шестое место рейтинга стран с самым доступным интернетом. URL: <https://www.championat.com/cybersport/news-4823271-rossiya-zanyala-shestoe-mesto-rejtinga-stran-s-samym-dostupnym-internetom.html> (дата обращения: 24.05.2023).

ровок для студентов в процессе дистанционного обучения;

- существующая система онлайн-приёма и реализация равенства в условиях перехода к гибридным технологиям не только в обучении, но и при приёме абитуриентов.

Данные обстоятельства, несмотря на кажущуюся давность событий, заставили авторов обратиться к исследованию последствий и условий этого перехода, отражённых во мнениях студентов по данному вопросу. В настоящее время сложно представить учебный процесс без цифровых технологий, больших данных, анализа цифрового следа. Они вошли в нашу жизнь, а во время пандемии, в период вынужденной самоизоляции наиболее наглядно продемонстрировали, как отдельные аспекты, в частности коммуникации, практически полностью перенесли в виртуальное пространство, формируя постоянно увеличивающийся массив цифровых следов из различных источников.

Научная проблема исследования заключается в противоречии между существующим объёмом неструктурированных данных цифровых следов студентов в социальных сетях и отсутствием научно-обоснованного и апробированного методологического подхода к анализу и оценке этих объёмных данных, что создаёт препятствия для фундаментального исследования взаимосвязи между активностью студентов в социальных сетях и их удовлетворённостью качеством образовательного процесса. Её значимость определяется тем, что анализ влияния обратной связи и поддержки на мотивацию и успеваемость студентов может помочь выявить важные связи и взаимное влияние между различными аспектами онлайн-образования и разработать эффективные стратегии и подходы к обучению, а также улучшить организацию программ для онлайн-образования. Для изучения указанных аспектов требуется проведение регулярного системного анализа с использованием научных методов и подходов, по-

зволяющих получить объективные данные и научно обоснованные выводы.

Практическая направленность данного исследования заключается в проведении анализа данных с применением инструментов больших данных. Полученные результаты и научно обоснованные выводы будут полезны для разработки инновационных стратегий и инструментов оценки и поддержки студентов.

Цель данного исследования – разработка научно обоснованного подхода к анализу больших объёмов данных цифровых следов студентов в социальных сетях для:

- а) эффективного выявления возникающих массовых проблем у студентов;
- б) поиска возможностей развития вузов и повышения качества образования за счёт автоматизированного анализа обсуждений.

Обзор литературы

Современное состояние исследований в области образовательной аналитики таково, что объём доступных данных растёт в геометрической прогрессии, но возможности для работы с ними не успевают формироваться и проходить апробацию с той же скоростью [7–9]. К такому выводу пришли авторы, изучив опубликованные результаты исследований применения анализа данных для решения образовательных проблем. В частности авторы данной работы отметили следующее.

Был выявлен ряд исследований, которые в интересах образовательной аналитики опираются на результаты опросных или административных данных с применением базовых методов анализа [10–13], выявляя так называемые маркеры самочувствия студентов. Однако существенным недостатком такого метода является ограниченность наборов данных, которые можно получить из указанных источников. Эти методы могут быть полезны, но они не всегда дают полную картину и не учитывают все аспекты студенческой активности и взаимодействия. Использование инструментов больших данных в образовательной аналитике предоставляет

значительные преимущества. Они позволяют обрабатывать и анализировать огромные объёмы данных, собранных из различных источников. Поэтому к этому инструменту обращается большое количество как российских, так и зарубежных учёных и практиков, работы которых изучили авторы.

Большинство исследователей сосредоточены на прогнозировании академической успеваемости [14–16]. Не отрицая важности подобных исследований, авторы однако считают, что для всестороннего повышения качества образования необходима такая же всесторонняя оценка учебного опыта студентов в ходе обучения. Актуальным направлением развития исследований в этой области является изучение учебного поведения учащихся и их уровня вовлечённости, взаимодействия с преподавателями, другими студентами, административными структурами вуза. Опубликованные работы, рассматривающие эту тему, представляют результаты анализа успеваемости и поведения, связанного с вовлечённостью, приложенными усилиями и обратной связью исключительно в цифровой образовательной среде [17–21].

В связи с тем, что подавляющее большинство студентов являются пользователями социальных сетей, при этом не ограничивая доступ в информации своего профиля [22–24], в последнее время растёт число исследований, которые используют цифровой след в социальных сетях (посты, подписки, лайки, фотографии со страниц профиля и на стене профиля) и алгоритмы машинного обучения для получения дополнительной информации о них. Так, например, показана возможность определить пол по сообщениям в *Twitter* [25], установить личностные характеристики пользователей на основе анализа их поведения в социальной сети [26]. Обнаружено, что интересы молодёжи дифференцируются в зависимости от их пола, возраста и успеваемости в обучении [27]. В работе [28] описывается разработка системы прогнозирования профессиональных интересов абитуриентов на основе данных социальных сетей.

Несмотря на разнородность подходов и методов в сфере образования, основанных на данных, авторы данной работы не встретили системных научных исследований, в которых социальные сети, содержащие информацию о предпочтениях и интересах пользователей, косвенно свидетельствуют об их коммуникативных и поведенческих особенностях, опыте и взаимосвязях с участниками различных сообществ [29], рассматривались как источник данных в образовательной аналитике.

Было обнаружено, что большинство исследователей используют аналитику обучения для улучшения удержания студентов; немногие из них сосредоточены на изучении удовлетворённости студентов в процессе обучения, возникающих у них проблемами [30]. Одним из наиболее известных направлений исследований в данной области является прогнозирующее моделирование [31]. Прогнозирующие модели приобрели популярность в образовании как конкурентная стратегия, которая выходит за рамки простого мониторинга успеваемости учащихся, позволяет реализовать упреждающее управление успехами студентов и раннюю разработку мер профилактического вмешательства для групп риска. Эта популярность подтверждается растущим объёмом исследований, посвящённых проблемам обучения с использованием прогнозирующих моделей.

Учитывая вышеописанный контекст, авторы сформировали гипотезу о том, что сообщения, оставляемые студентами в социальных сетях, представляют собой ценный источник данных для образовательной аналитики: на основании их анализа можно не только выявить общие паттерны настроения, но и отследить их для отдельных групп университетов, сравнивая графики изменения эмоциональной окраски сообщений по тематическим направлениям, выделить возможности улучшения, развития, болевые точки в образовательном процессе и его сопровождении.

Методология, материалы и методы

Работа основана на продолжающемся тематическом проекте изучения цифровых следов студентов российских вузов в социальных сетях, которое является частью более широкого исследования, содержащего аспекты программного обеспечения, данных и алгоритмов, особенно в части инструментария анализа больших данных.

Разработанная и апробированная методика анализа цифровых следов студентов российских вузов в социальных сетях включает следующие этапы.

1. Сбор данных. Для анализа использовались дата-сети сообщений и комментариев из социальных сетей, содержащие около 3 миллионов постов и комментариев. Данные были собраны из 9187 официальных и неофициальных пабликов 679 высших учебных заведений России, включая филиалы. Период накопления данных, выбранных для анализа, – с сентября 2021 г. по июнь 2022 г., что определяется сроками начала и окончания учебного года в подавляющем большинстве российских вузов.

2. Использование информационно-аналитической платформы (ИАП). В качестве инструмента анализа была задействована ИАП *PolyAnalyst*. Эта платформа предоставляет широкие возможности автоматизации работы с большими данными. В рамках данного исследования использовались функции извлечения тональностей, ключевых слов, сущностей и отношений из неструктурированных текстов, классификации и кластеризации документов, а также визуализации корреляций между темами и атрибутами.

3. Анализ тональности. Проводился анализ относительного количества постов и комментариев по эмоциональному настрою. Использовалось отношение количества сообщений по теме к общему количеству выгруженных сообщений. Для извлечения тональностей в ИАП *PolyAnalyst* был разработан и настроен соответствующий сценарий. Часть данных (10%) была проверена

вручную для определения валидности результатов. Ручная выверка производилась силами 35 разметчиков – студентов вузов, прошедших ознакомительное обучение. Каждый элемент был проверен дважды двумя разными разметчиками для обеспечения целостности данных.

4. Кластерный анализ. После анализа тональности проводился кластерный анализ сообщений из социальной сети на основе ключевых слов. Соответствующие сценарии были настроены в ИАП *PolyAnalyst*. Это позволило структурировать данные и визуализировать тренды в обсуждениях для лучшего понимания всего массива данных и его структуры.

5. Анализ настроений пользователей. Была проанализирована динамика настроений пользователей социальной сети «ВКонтакте» по вопросам взаимодействия преподавателей / научных руководителей и студентов/аспирантов в разрезе категорий вузов (научно-исследовательские, опорные, федеральные и прочие). Это позволило определить, влияет ли принадлежность вуза к одной из категорий на удовлетворённость студентов качеством предоставляемого образования.

Паттерны для выгрузки данных обоснованы целью исследования и определены на основе ключевых слов или хэштегов, связанных с этими темами, которые также стали результатом авторской разработки лингвистических маркеров – критериев тональности сообщений в социальных сетях. Их разработка включала ознакомление с данными, изучение их структуры и содержания, извлечение тональностей, ключевых слов, сущностей и отношений. Методы понимающего восприятия текста основаны на алгоритмах обработки естественного языка (*Natural Language Processing, NLP*), которые позволяют автоматически выделить смысловую информацию из текстов. Концепция паттернов для выгрузки данных из социальных сетей является частью области исследований, связанных с анализом текстов и обработкой

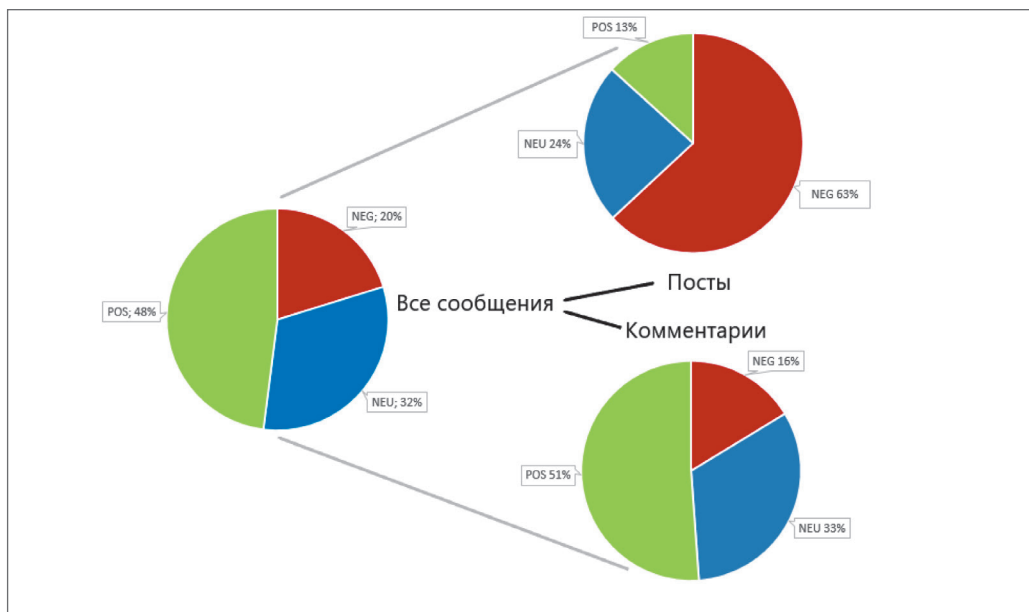


Рис. 1. Распределение настроений в категории «Особенности взаимодействия преподавателей и студентов»

Fig. 1. Distribution of sentiments in the category “ Peculiarities of interaction between teachers and students”

естественного языка. Процесс их формирования также включал несколько этапов:

1) изучение структуры сообщений, специфики языка, используемого пользователями, и других факторов, которые могут повлиять на процесс анализа;

2) сбор обучающих данных путём аннотации сообщений экспертами (25%) и автоматических методов классификации ИАП *PolyAnalyst* (75%);

3) определение ключевых слов и хэштегов на основе анализа обучающих данных;

4) применение методов обработки естественного языка методами анализа естественного языка: классификация текстов, выделение ключевых слов, распознавание именованных сущностей и анализ синтаксиса;

5) на основе обучающих данных и результатов анализа применение методов машинного обучения или статистических подходов для разработки паттернов (паттерны были протестированы на массиве 103 тысяч сообщений, чтобы убедиться в их эффективности и точности).

Использование паттернов для выгрузки данных из социальных сетей является активной областью исследований и в настоящее время не существует единых универсальных подходов и методов, которые могут быть применены в этом процессе, что и привело к необходимости разработки авторского набора в интересах исследования.

Результаты

1. Выборка сообщений по категории «Особенности взаимодействия преподавателей и студентов» самая многочисленная – 41,72% от общего количества анализируемых сообщений. Распределение между негативными, позитивными и нейтральными сообщениями по категории представлено на *рисунке 1*.

2. Тематический sentiment-анализ по разработанным паттернам выявил отдельные тематические подкатегории в сообщениях пользователей университетских сообществ в социальной сети «ВКонтакте» (*табл. 1*). Для измерения надёжности опре-

Таблица 1

Отдельные тематические подкатегории в сообщениях

Table 1

Separate thematic subcategories in messages

Тематическая подкатегория	Описание	Каппа Коэна	Процентное согласие	Частота, %
Одобрение дистанционного обучения и применяемых технологий, удовлетворённость взаимодействием с преподавателями	Сообщения (посты, комментарии) поддерживают дистанционное образование либо (1) в любом случае, либо (2) потому, что нет другого / лучшего варианта, чем дистанционное образование, как например, во время кризиса COVID-19.	0,87	97,1	19
Негативное отношение к дистанционному обучению	В сообщениях (постах, комментариях) говорится, что дистанционное образование уступает очному обучению, перечисляются аргументы в пользу того, что оно неэффективно.	0,845	92,6	16
Низкий уровень технологической и организационной готовности вузов к реализации образовательного процесса и сервисов для студентов дистанционно	В сообщениях (постах, комментариях) говорится, что вузы не имеют достаточных технических возможностей / инфраструктуры для реализации дистанционного обучения, а студенты сталкиваются с трудностями там, где, по их мнению, этого можно было избежать при должной готовности.	0,798	99,1	14
Неравные возможности при дистанционном обучении	Сообщения (посты, комментарии) о каком-либо неравенстве среди студентов с позиции доступа к дистанционному образованию (например, наличие компьютеров, качество подключения к Интернету, отсутствие в общежитии или дома комфортных условий для онлайн-обучения, отсутствие персонального рабочего места и т. п.).	0,912	95,3	12
Неготовность или нежелание преподавателей осуществлять обучение в дистанционном формате, использовать современные технологии	Сообщения (посты, комментарии) ссылаются на низкий уровень готовности преподавателей использовать интерактивные технологии в обучении, нежелание общаться посредством камеры, микрофона, чата и других средств связи, недостаточное качество электронных контентов, проблемы с планированием времени и расписанием.	0,792	98,2	6
Аттестационные мероприятия	Сообщения (посты, комментарии), в которых студенты излагают беспокойство по поводу доступности, справедливости аттестации в формате дистанционного обучения.	0,81	96,2	6
Недовольство работой административно-управленческого персонала вузов	Сообщения (посты, комментарии), в которых пишут о том, что административно-управленческий персонал вузов неспособен справиться с ситуацией полностью или частично, а также с информированием, поддержкой студентов, своевременным принятием необходимых мер.	0,852	99,0	5

Продолжение Таблицы 1

Тематическая подкатегория	Описание	Каппа Коэна	Процентное согласие	Частота, %
Проблемы физического и психического здоровья	В сообщениях (постах, комментариях) говорится, что у пользователей возникают проблемы с физическим и психическим здоровьем, которые в свою очередь вызваны особенностями дистанционного обучения.	0,821	94,9	4
Социальная неудовлетворённость студентов при дистанционном обучении	Сообщения (посты, комментарии), несущие недовольство последствиями дистанционного или гибридного обучения для социальной жизни студентов, замечания о нехватке взаимодействия с одногруппниками и/или фасилитирующего воздействия со стороны преподавателя, поддержки.	0,61	95,6	4
Эмоциональные выражения	Сообщения (посты, комментарии) содержат явное выражение эмоций по поводу дистанционного обучения без логического обоснования причин («бесит», «раздражает», «я кайфую» и т. п.).	0,69	98,53	6
Юмор, сарказм	Сообщения (посты, комментарии), содержащие шутки по поводу взаимодействия с преподавателями в процессе дистанционного обучения.	0,771	94,15	8

деления частоты для качественных (категориальных) позиций авторами применялась Каппа Коэна как более надёжная мера, чем простой расчёт процентного согласия, поскольку Каппа Коэна учитывает возможность случайного совпадения. Однако авторы рассчитали и процентное согласие также для каждой подкатегории, чтобы удостовериться в точности полученных результатов.

3. Поддержка дистанционного (или гибридного) обучения составила в исследуемом периоде 19%, что является высоким показателем, несмотря на значительно снизившийся в обществе уровень тревожности в связи с пандемией COVID-19. Рассмотрев относительное количество постов и комментариев по эмоциональному настрою к общему количеству опубликованных сообщений и наложив их на график заболеваемости COVID-19, авторы получили результаты, представленные на *рисунке 2*.

4. Были выделены наиболее обсуждаемые проблемные вопросы, касающиеся обучения:

- недовольство студентов работой администрации вузов в отдельных аспектах, таких как доступность получения привычных сервисов, например, справок об обучении, консультаций по организации обучения, доступа к расписанию – 9,18%;
 - неготовность или нежелание преподавателей осуществлять обучение в дистанционном формате, использовать современные технологии – 4,67%;
 - социальная неудовлетворённость студентов при дистанционном обучении – 3,95%.
5. Были выделены наиболее обсуждаемые проблемные вопросы, касающиеся здоровья студентов:
- проблемы физического и психического здоровья, которые возникают у студентов при смене режима обучения после очного, однако они могут быть нивелированы при соблюдении режимов работы и отдыха – 14,35%;
 - проблемы демографического характера (упоминания об ограничениях доступа к

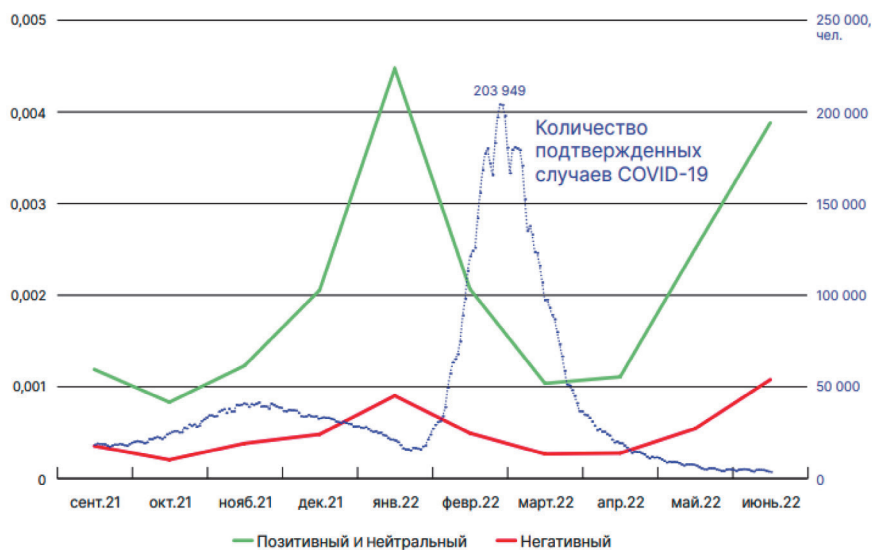


Рис. 2. Относительное количество постов и комментариев категории «Особенности взаимодействия преподавателей/ научных руководителей и студентов/ аспирантов»

Fig. 2. Relative number of posts and comments in the category "Peculiarities of interaction between teachers/ supervisors and students/ graduate students"

средствам дистанционного обучения и технологической инфраструктуре у студентов из семей низкого социально-экономического статуса, в редких случаях – лиц с ограниченными возможностями, а также тех, у кого нет доступа к надёжному интернет-соединению или устройствам, необходимым для участия в онлайн-образовании из-за работы вахтовым методом или проживания в отдалённых районах) – 2,32%.

6. Студенты после опыта пандемии и вынужденного перехода на дистанционное обучение, а затем – возврата к очному обучению высказывают совершенно чёткие предпочтения в отношении дистанта. В частности, наиболее часто встречаются положительные упоминания:

- удобство и интуитивная понятность курсов в электронных средах обучения – 26,1%;
- использование информационно-коммуникационных технологий в учебном процессе в целом – 22,23%;
- использование дискуссионных технологий и форумов онлайн – 12,82%.

7. В таблице 2 представлены результаты изучения данных по академической поддержке студентов при гибридном и дистанционном обучении в исследуемый период.

Чтобы выразить своё восприятие контекста обучения, студенты ссылаются на организационные проблемы и делятся просьбами о более продуманной организации обучения и заданий (рис. 3).

10. Рассматривая динамику настроений пользователей соцсети «ВКонтакте» по вопросам взаимодействия преподавателей/ научных руководителей и студентов/ аспирантов в разрезе категорий вузов (рис. 4, 5), авторы обратили внимание, что кривые для опорных и федеральных университетов более гладкие для всех эмоциональных окрасок, что говорит о большей устойчивости данных типов образовательных организаций в плане обеспечения взаимодействия между преподавателями и студентами.

Сравнив данные результаты с исследованиями по смежным тематикам, авторы отметили, что полученные ими результаты с фиксированными эффектами свидетельствуют

Таблица 2

Результаты исследования данных по удовлетворённости студентов академической поддержкой

Table 2

Findings from a study of data on student satisfaction with academic support

Условия для онлайн-обучения, созданные вузом		Упоминание среди релевантных сообщений, %
Негативное отношение		24,5
Нейтральное отношение		22,3
Позитивное отношение		53,2
Всего		100,0
Наличие доступа к библиотечным ресурсам		Упоминание среди релевантных сообщений, %
Говорят о наличии доступа в любом контексте	Негативное отношение	18,1
	Нейтральное отношение	60,2
	Позитивное отношение	19,5
Говорят об отсутствии доступа		2,2
Всего		100
Наличие доступа к лабораторному оборудованию или виртуальным лабораторным работам		Упоминание среди релевантных сообщений, %
Говорят о наличии доступа в любом контексте	Негативное отношение	11,2
	Нейтральное отношение	62,1
	Позитивное отношение	12,9
Говорят об отсутствии доступа		13,8
Всего		100
Доступ к программному обеспечению, необходимому в ходе обучения		Упоминание среди релевантных сообщений, %
Говорят о наличии в любом контексте	Негативное отношение	12,1
	Нейтральное отношение	21,5
	Позитивное отношение	56,6
Говорят об отсутствии доступа		9,8
Всего		100

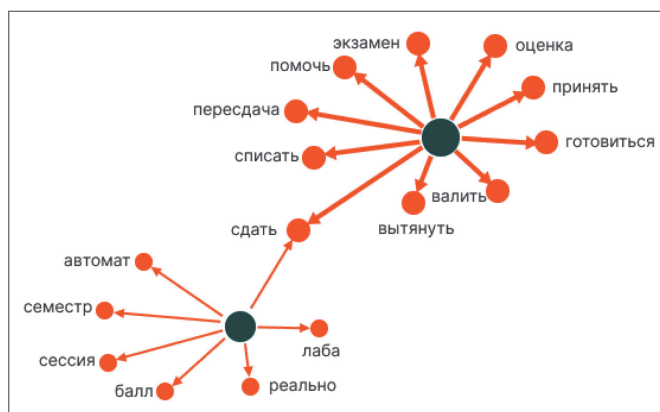


Рис. 3. Кластерный анализ организации обучения
 Fig. 3. Cluster analysis of the organization of education

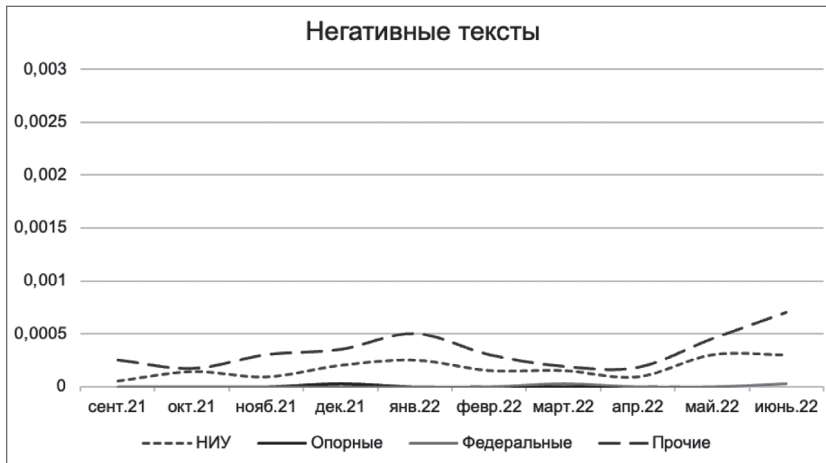


Рис. 4. Динамика негативных настроений пользователей социальной сети «ВКонтакте» по вопросам взаимодействия преподавателей и студентов в разрезе категорий вузов

Fig. 4. Dynamics of negative moods of users of the social network VKontakte on the issues of interaction between teachers and students in the context of university categories

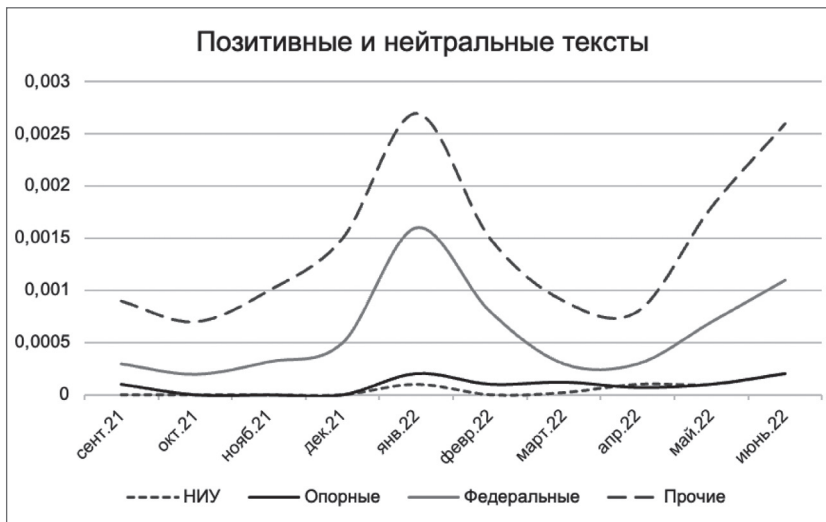


Рис. 5. Динамика позитивных и нейтральных настроений пользователей социальной сети «ВКонтакте» по вопросам взаимодействия преподавателей и студентов в разрезе категорий вузов

Fig. 5. Dynamics of positive and neutral moods of users of the social network VKontakte on the issues of interaction between teachers and students in the context of university categories

о существенном улучшении удовлетворённости студентов качеством взаимодействия с представителями вузов, в частности преподавателями. Увеличение частоты социальных контактов, уменьшение беспокойства о будущем в значительной степени способствовали повышению удовлетворённости (на

21,03% по сравнению с тем же периодом в прошлом году).

Выводы

Разработан научно обоснованный подход к анализу больших объёмов данных цифровых следов студентов в социальных сетях,

который позволяет эффективно выявлять возникающие и наиболее обсуждаемые проблемы у студентов, а также выделять болевые точки, дающие возможности роста, развития вузов и улучшения характеристик образовательного процесса, сопровождения студентов и т. д.

Полученные результаты показывают:

- пики активности в публикациях студентов в социальных сетях связаны с окончанием семестра (январь и июнь) – это традиционная картина для учебного процесса, так как в это время студенты сдают сессию или защищают выпускные квалификационные работы;

- сравнивая по представленной методике массив сообщений отдельно взятого вуза, группы вузов, объединённых по какому-либо признаку, с общим массивом цифровых следов, можно обратить внимание на периоды аномальной активности в обсуждениях студентов в связке с эмоциональной окраской сообщений; этот механизм можно использовать для оперативного реагирования, тактических мероприятий и разработки стратегий развития, для выявления проблем и возможностей студентов инструментами больших данных.

Исследование свидетельствует в пользу тщательного и регулярного изучения цифрового следа студентов не только в любой образовательной среде, но и в других, отвлечённых средах – выяснения их мнений, «выслушивания» их голосов, реакций, размышлений и действий. Это особенно важно в условиях, когда живого общения мало, а традиционные процессы и процедуры нарушены, например, при экстренном дистанционном обучении, равно как и возвращении после него к очному обучению, но в равной степени рекомендуется в ситуациях, не связанных с чрезвычайными мерами.

Обсуждение

Дистанционный формат взаимодействия выявил множество проблем, с которыми участники образовательного процесса ещё

не сталкивались. Пандемия показала настоящий уровень цифровой трансформации российских университетов, уровень готовности к дистанционному формату взаимодействия с участниками образовательного процесса, уровень автоматизации административных процессов. Данное исследование направлено на решение задачи применения технологий больших данных, машинного обучения и искусственного интеллекта, анализ неструктурированных данных из социальных сетей и вносит вклад в развитие образовательной аналитики благодаря тому, что представленные методы позволяют получить ценные инсайты и представления о взаимодействии между участниками образовательного процесса. Анализ цифровых следов студентов может помочь в понимании и оценке качества образования, выявлении трендов и проблем, а также определении потенциала для улучшения образовательного процесса. Использование информационно-аналитической платформы, в данном случае – ИАП *PolyAnalyst*, предоставляет возможности автоматизации анализа больших данных, включая извлечение тональностей, ключевых слов, сущностей, классификацию и кластеризацию данных. Это позволило авторам структурировать и визуализировать данные для лучшего понимания их содержания и трендов. Использование ИАП *PolyAnalyst* и применение кластеризации позволяют проводить сложный анализ больших объёмов данных и визуализировать результаты, что способствует развитию и автоматизации аналитических процессов. Следующие шаги в этом направлении могут быть нацелены на автоматизацию обработки данных социальных сетей в режиме реального времени.

Однако необходимо подчеркнуть, что предлагаемая методология имеет свои ограничения в качестве концептуальной дорожной карты для дальнейших исследований. Самое главное из них заключается в том, что выборка сообщений производится в заранее отобранных для этого сообществах, в данном случае – сообществах вузов в соци-

альной сети «ВКонтакте». То есть заведомо ограничили область поиска. Второе ограничение исследования состоит в том, что выбранные инструменты не учитывают послания, представленные в графическом формате («смайлики», картинки). Наконец, третье ограничение – риск приостановки доступа к данным, который определяется политикой работы с данными социальных сетей.

Заключение

В исследовании изучены социальные и организационные аспекты онлайн-образования через анализ данных сообщений в социальной сети «ВКонтакте» инструментами больших данных. Это помогло выявить важные связи и влияния между различными аспектами онлайн-образования.

Авторы акцентировали внимание не только на проблемах и ограничениях, которые появились в университетах по всему миру, но и возможностях, которые открылись как со стороны обучающихся, так и со стороны обучающихся, на управленческих решениях, которые помогут повысить доступность и качество высшего образования.

Благодаря анализу больших данных можно видеть, что студенты активно участвуют в структурировании своей жизни и образовательного опыта, принимая на себя важную проблематизирующую роль в переосмыслении того, каким должно быть их собственное обучение и образование в целом. Авторы смогли подтвердить гипотезу исследования о том, что на основании анализа цифровых следов студентов в социальных сетях можно не только выявить общие паттерны настроения, но и отследить их для отдельных групп университетов, сравнивая графики изменения эмоциональной окраски сообщений по тематическим направлениям, выделить возможности улучшения, развития, болевые точки в образовательном процессе и его сопровождении.

Проблемы, обозначенные в статье, характерны не только для аналитики дистанционного обучения. Чтобы продвинуть образо-

вательную аналитику вперёд, необходимо использовать междисциплинарный подход. Решающее значение для этого будут иметь постоянная проверка и подтверждение достоверности собираемых больших данных, целостности методов их обработки. Образовательная аналитика проблем, возникающих у студентов, а также их перспективных возможностей на основе больших данных может рассматриваться как новая форма оценки качества образования в цифровую эпоху, которая будет играть всё более важную роль в образовательных исследованиях и практиках.

Список литературы

1. *Marginson S.* Imagining the global // Handbook of globalization and higher education: ed. by R. King, S. Marginson, R. Naidoo. 2011. P. 10–39. URL: https://www.researchgate.net/publication/285738663_Imagining_the_global (дата обращения: 20.04.2023).
2. *Chapleo C., O'Sullivan H.* Contemporary thought in higher education marketing // Journal of Marketing for Higher Education. 2017. Vol. 27. No. 2. P. 159–161. DOI: 10.1080/08841241.2017.1406255
3. *Горбунова Е.В.* Выбытия студентов из вузов: исследования в России и США // Вопросы образования. 2018. № 1. С. 110–131. DOI: 10.17323/1814-9545-2018-1-110-131
4. *Груздев И.А., Горбунова Е.В., Фруммин И.Д.* Студенческий отсев в российских вузах: к постановке проблемы // Вопросы образования. 2013. №2 (октябрь). С. 67–81. DOI: 10.17323/1814-9545-2013-2-67-81
5. *Arria A.M., Garnier-Dykstra L.M., Caldeira K.M., Vincent K.B., Winick E.R., O'Grady K.E.* Drug use patterns and continuous enrollment in college: results from a longitudinal study // Journal of Studies on Alcohol and Drugs. 2013. Vol. 74. No. 1. P. 71–83. DOI: 10.15288/jsad.2013.74.71
6. *Дзюбан В.В.* Проблема внедрения цифровых технологий в систему образования в XX–XXI в. // Архонт. 2021. № 6 (27). С. 34–39. EDN: GHGFYH.
7. *Vieira C., Parsons P., Byrd V.* Visual learning analytics of educational data: A systematic literature review and research agenda // Computers & Education. 2018. No. 122. P. 119–135. DOI: 10.1016/j.compedu.2018.03.018

8. *Huda M., Maseleno A., Shabrill M., Jasmi K.A., Mustari I., Basiron B.* Exploring adaptive teaching competencies in big data era // International Journal of Emerging Technologies in Learning (ijET). 2017. Vol. 12. No. 03. P. 68–83. DOI: 10.3991/ijet.v12i03.6434
9. *Imani M., Montazer G.A.* A survey of emotion recognition methods with emphasis on e-learning environments // Journal of Network and Computer Applications. 2019. No. 147. Article no. 102423. DOI: 10.1016/j.jnca.2019.102423
10. *Sultan P., Wong H.Y.* Service quality in a higher education context: An integrated model // Asia Pacific Journal of Marketing and Logistics. 2012. Vol. 24. No. 5. P. 755–784. DOI: 10.1108/13555851211278196
11. *Wulff A.* Global Education Governance in the Context of COVID-19: Tensions and Threats to Education as a Public Good // Development. 2021. No. 64. P. 74–81. DOI: 10.1057/s41301-021-00293-1
12. *Raju D., Schumacker R.* Exploring student characteristics of retention that lead to graduation in higher education using data mining models // Journal of College Student Retention: Research, Theory & Practice. 2015. No. 16. P. 563–591. DOI: 10.2190/CS.16.4.e
13. *Perez B., Castellanos C., Correal D.* Applying data mining techniques to predict student dropout: a case study // 2018 IEEE 1st Colombian Conference on Applications in Computational Intelligence (ColCACI). 2018. P. 1–6. DOI: 10.1109/ColCACI.2018.8484847
14. *Baek C., Doleck T.* Educational data mining versus learning analytics: A review of publications from 2015 to 2019 // Interactive Learning Environments. 2021. P. 1–23. DOI: 10.1080/10494820.2021.1943689
15. *de Oliveira C.F., Sobral S.R., Ferreira M.J., Moreira F.* How does learning analytics contribute to prevent students dropout in higher education: A systematic literature review // Big Data and Cognitive Computing. 2021. Vol. 5 (4). No. 64. P. 1–33. DOI: 10.3390/bdcc5040064
16. *Larrabee Sønderlund A., Hughes E., Smith J.* The efficacy of learning analytics interventions in higher education: A systematic review // British Journal of Educational Technology. 2019. Vol. 50. No. 5. P. 2594–2618. DOI: 10.1111/bjet.12720
17. *Rastrullo-Guerrero J.L., Gómez-Pulido J.A., Durán-Domínguez A.* Analyzing and Predicting students performance by means of machine learning: A review // Applied Sciences. 2020. Vol. 10. No. 3. Article no. 1042. P. 1–16. DOI: 10.3390/app10031042
18. *Goel Y., Goyal R.* On the effectiveness of self-training in MOOC drop-out prediction // Open Computer Science. 2020. Vol. 10. No. 1. P. 246–258. DOI: 10.1515/comp-2020-0153
19. *Guo S., Zeng D., Dong S.* Pedagogical data analysis via federated learning toward education 4.0 // American Journal of Education and Information Technology. 2020. Vol. 4. No. 2. P. 55–56. DOI: 10.11648/j.ajeit.20200402.13
20. *Hasan R., Palaniappan S., Mahmood S., Abbas A., Sarker K.U., Sattar M.U.* Predicting student performance in higher educational institutions using video learning analytics and data mining techniques // Applied Sciences (Switzerland). 2020. Vol. 10. No. 11. Article no. 3894. DOI: 10.3390/app10113894
21. *Mai T.T., Bezbradica M., Crane M.* Learning behaviours data in programming education: Community analysis and outcome prediction with cleaned data // Future Generation Computer Systems. 2022. No. 127. P. 42–55. DOI: 10.1016/j.future.2021.08.026
22. *Von Hippel P.T., Hofflinger A.* The data revolution comes to higher education: identifying students at risk of dropout in Chile // Journal of Higher Education Policy and Management. 2020. 24 p. DOI: 10.1080/1360080X.2020.1739800
23. *Cannistrà M., Masci C., Ieva F., Agasisti, T., Paganoni A.M.* Not the magic algorithm: modelling and early-predicting students dropout through machine learning and multilevel approach // MOX-Modelling and Scientific Computing, Department of Mathematics, Politecnico di Milano, Via Bonardi (Milan). 2020. No. 41. 33 p. URL: <https://www.mate.polimi.it/biblioteca/add/qmox/41-2020.pdf> (дата обращения: 25.04.2023).
24. *Azcona D., Hsiao I.H., Smeaton A.F.* Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints // User Modeling and User-Adapted Interaction. 2019. No. 29. P. 759–788. DOI: 10.1007/s11257-019-09234-7
25. *Miller Z., Dickinson B., Hu W.* Gender prediction on Twitter using stream algorithms with n-gram character features // International Journal of Intelligence Science. 2012. No. 2. P. 143–148. DOI: 10.4236/ijis.2012.224019

26. Поливанова К.Н., Смирнов И.Б. Что в профиле тебе моем: Данные «ВКонтакте» как инструмент изучения интересов современных подростков // Вопросы образования. 2017. No. 2. С. 134–152. DOI: 10.17323/1814-9545-2017-2-134-152
27. Gosling S.D., Augustine A., Vazire S., Holtzman N., Gaddis S. Manifestations of personality in online social networks // *Cyberpsychology, Behavior, and Social Networking*. 2011. Vol. 14. No. 9. P. 483–488. DOI: 10.1089/cyber.2010.0087
28. Степаненко А.А., Резанова З.И., Гойко В.А. Автоматическая классификация контента персональных страниц пользователей социальной сети «ВКонтакте» как маркеров профессиональных интересов абитуриента // *Гуманитарная информатика*. 2018. No. 15. С. 20–26. DOI: 10.17223/23046082/15/1
29. Dascalu M., Popescu E., Becheru A., Crossley S., Trausan-Matu S. Predicting academic performance based on students' blog and microblog posts // *European Conference on Technology Enhanced Learning*. 2016. P. 370–376. DOI: 10.1007/978-3-319-45153-4_29
30. Hernández-de-Menéndez M., Morales-Méndez R., Escobar C.A., Ramírez Mendoza R.A. Learning analytics: state of the art // *International Journal on Interactive Design and Manufacturing*. 2022. No. 16. P. 1209–1230. DOI: 10.1007/s12008-022-00930-0
31. Brooks C.A., Thompson C.D.S. Chapter 5: Predictive Modelling in Teaching and Learning. 2017. URL: <https://www.semanticscholar.org/paper/Chapter-5-%3A-Predictive-Modelling-in-Teaching-and-BrooksThompson/2cd4901b07f3562f98e1e56dc5712e8bc03bdc2e> (дата обращения: 20.04.2023).

Благодарности. Исследование выполнено при поддержке Программы развития ТГУ («Приоритет-2030»). Мы также выражаем искреннюю благодарность рецензентам журнала за внимание, уделенное нашей статье.

Статья поступила в редакцию 05.07.2023

Принята к публикации 22.08.2023

References

1. Marginson, S. (2011). Imagining the Global. In: R. King, S. Marginson, R. Naidoo (Eds.). *Handbook of Globalization and Higher Education*. Pp. 10-39. Available at: https://www.researchgate.net/publication/285738663_Imagining_the_global (accessed 04.20.2023).
2. Chapleo, C., O'Sullivan, H. Contemporary Thought in Higher Education Marketing. (2017). *Journal of Marketing for Higher Education*. Vol. 27, no. 2, pp. 159-161, doi: 10.1080/08841241.2017.1406255
3. Gorbunova, E.V. (2018). Students Leaving Universities: Studies in Russia and the USA. *Voprosy obrazovaniya = Education Studies*. No. 1, pp. 110-131, doi: 10.17323/1814-9545-2018-1-110-131
4. Gruzdev, I.A., Gorbunova, E.V., Frumin, I.D. (2013). Student Dropout in Russian Universities: Towards the Formulation of the Problem. *Voprosy obrazovaniya = Education Studies*. No. 2 (October), pp. 67-81, doi: 10.17323/1814-9545-2013-2-67-81 (In Russ., abstract in Eng.).
5. Arria, A.M., Garnier-Dykstra, L.M., Caldeira, K.M., Vincent, K.B., Winick, E.R., O'Grady, K.E. (2013). Drug Use Patterns and Continuous Enrollment in College: Results from a Longitudinal Study. *Journal of Studies on Alcohol and Drugs*. Vol. 74, no. 1, pp. 71-83, doi: 10.15288/jasad.2013.74.71
6. Dzyuban, V.V. (2021). The Problem of Introducing Digital Technologies into the Education System in the XX-XXI Centuries. *Arkhont [Archon]*. No. 6 (27), pp. 34-39. Available at: https://elibrary.ru/download/elibrary_48021742_17885110.pdf (accessed 04.20.2023). (In Russ., abstract in Eng.).
7. Vieira, C., Parsons, P., Byrd, V. (2018). Visual Learning Analytics of Educational Data: A Systematic Literature Review and Research Agenda. *Computers & Education*. No. 122, p. 119-135, doi: 10.1016/j.compedu.2018.03.018

8. Huda, M., Maselena, A., Shahrill, M., Jasmi, K. A., Mustari, I., Basiron, B. (2017). Exploring Adaptive Teaching Competencies in Big Data Era. *International Journal of Emerging Technologies in Learning (ijET)*. Vol. 12, no. 03, pp. 68-83, doi: 10.3991/ijet.v12i03.6434
9. Imani, M., Montazer, G.A. (2019). A Survey of Emotion Recognition Methods with Emphasis on E-Learning Environments. *Journal of Network and Computer Applications*. No. 147, article no. 102423, doi: 10.1016/j.jnca.2019.102423
10. Sultan, P., Wong, H.Y. (2012). Service Quality in a Higher Education Context: An Integrated Model. *Asia Pacific Journal of Marketing and Logistics*. Vol. 24, no. 5, pp. 755-784, doi: 10.1108/13555851211278196
11. Wulff, A. (2021). Global Education Governance in the Context of COVID-19: Tensions and Threats to Education as a Public Good. *Development*. No. 64, pp. 74-81, doi: 10.1057/s41301-021-00293-1
12. Raju, D., Schumacker, R. (2015). Exploring Student Characteristics of Retention That Lead to Graduation in Higher Education Using Data Mining Models. *Journal of College Student Retention: Research, Theory & Practice*. No. 16, pp. 563-591, doi: 10.2190/CS.16.4.e
13. Perez, B., Castellanos, C., Correal, D. (2018). Applying Data Mining Techniques to Predict Student Dropout: A Case Study. In: *2018 IEEE 1st Colombian Conference on Applications in Computational Intelligence (ColCACI)*. Pp. 1-6, doi: 10.1109/ColCACI.2018.8484847
14. Baek, C., Doleck, T. (2021). Educational Data Mining Versus Learning Analytics: A Review of Publications from 2015 to 2019. *Interactive Learning Environments*. Pp. 1-23, doi: 10.1080/10494820.2021.1943689
15. de Oliveira, C.F., Sobral S.R., Ferreira M.J., Moreira F. (2021). How Does Learning Analytics Contribute to Prevent Students Dropout in Higher Education: A Systematic Literature Review. *Big Data and Cognitive Computing*. Vol. 5 (4), no. 64, pp. 1-33, doi: 10.3390/bdcc5040064
16. Larrabee Sønderlund, A., Hughes, E., Smith, J. (2019). The Effectiveness of Learning Analytics Interventions in Higher Education: A Systematic Review. *British Journal of Educational Technology*. Vol. 50, no. 5, pp. 2594-2618, doi: 10.1111/bjet.12720
17. Rastrollo-Guerrero, J.L., Gómez-Pulido, J.A., Durán-Domínguez, A. (2020). Analyzing and Predicting Student Performance by Means of Machine Learning: A Review. *Applied Sciences*. Vol. 10, no. 3, article no. 1042, pp. 1-16, doi: 10.3390/app10031042
18. Goel, Y., Goyal, R. (2020). On The Effectiveness of Self-Training in MOOC Drop-Out Prediction. *Open Computer Science*. Vol. 10, no. 1, pp. 246-258, doi: 10.1515/comp-2020-0153
19. Guo, S., Zeng, D., Dong, S. (2020). Pedagogical Data Analysis Via Federated Learning Toward Education 4.0. *American Journal of Education and Information Technology*. Vol. 4, no. 2, pp. 55-56, doi: 10.11648/j.ajeit.20200402.13
20. Hasan, R., Palaniappan, S., Mahmood, S., Abbas, A., Sarker, K.U., Sattar, M.U. (2020). Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques. *Applied Sciences (Switzerland)*. Vol. 10, no. 11, article no. 3894, doi: 10.3390/app10113894
21. Mai, T.T., Bezbradica, M., Crane, M. (2022). Learning Behaviors Data in Programming Education: Community Analysis and Outcome Prediction with Cleaned Data. *Future Generation Computer Systems*. No. 127, pp. 42-55, doi: 10.1016/j.future.2021.08.026
22. Von Hippel, P.T., Hofflinger, A. (2020). The Data Revolution Comes to Higher Education: Identifying Students at Risk Of Dropout in Chile. *Journal of Higher Education Policy and Management*. 24 p., doi: 10.1080/1360080X.2020.1739800
23. Cannistrà, M., Masci, C., Ieva, F., Agasisti, T., Paganoni, A.M. (2020). Not the magic algorithm: modeling and early-predicting students dropout through machine learning and multilevel ap-

- proach. In: *MOX-Modelling and Scientific Computing, Department of Mathematics, Politecnico di Milano, Via Bonardi (Milan)*. No. 41. 33 p. Available at: <https://www.mate.polimi.it/biblioteca/add/qmox/41-2020.pdf> (accessed 04.20.2023).
24. Azcona, D., Hsiao, I.H., Smeaton, A.F. (2019). Detecting Students-At-Risk in Computer Programming Classes with Learning Analytics from Students' Digital Footprints. *User Modeling and User-Adapted Interaction*. No. 29, pp. 759-788, doi: 10.1007/s11257-019-09234-7
 25. Miller, Z., Dickinson, B., Hu, W. (2012). Gender Prediction on Twitter Using Stream Algorithms With N-Gram Character Features. *International Journal of Intelligence Science*. No. 2, pp. 143-148, doi: 10.4236/ijis.2012.224019
 26. Polivanova, K.N., Smirnov, I.B. (2017). What's in My Profile for You: Vkontakte Data as a Tool for Studying the Interests of Modern Teenagers. *Voprosy obrazovaniya = Education Studies*. No. 2, pp. 134-152, doi: 10.17323/1814-9545-2017-2-134-152
 27. Gosling, S.D., Augustine, A., Vazire, S., Holtzman, N., Gaddis, S. (2011). Manifestations of Personality in Online Social Networks. *Cyberpsychology, Behavior, and Social Networking*. Vol. 14, no. 9, pp. 483-488, doi: 10.1089/cyber.2010.0087
 28. Stepanenko, A.A., Rezanova, Z.I., Goiko, V.L. (2018). Automatic Classification of the Content of Personal Pages of Users of the Social Network "Vkontakte" as Markers of the Professional Interests of an Applicant. *Humanitarian Informatics*. No. 15, pp. 20-26, doi:10.17223/23046082/15/1
 29. Dascalu, M., Popescu, E., Becheru, A., Crossley, S., Trausan-Matu, S. (2016). Predicting Academic Performance Based on Students' Blog and Microblog Posts. *European Conference on Technology Enhanced Learning*. Pp. 370-376, doi: 10.1007/978-3-319-45153-4_29
 30. Hernández-de-Menéndez, M., Morales-Menéndez, R., Escobar, C.A., Ramírez Mendoza, R.A. (2022). Learning analytics: state of the art. *International Journal on Interactive Design and Manufacturing*. No. 16, pp. 1209-1230, doi: 10.1007/s12008-022-00930-0
 31. Brooks, C.A., Thompson, C.D.S. (2017). Chapter 5: Predictive Modeling in Teaching and Learning. Available at: <https://www.semanticscholar.org/paper/Chapter-5-%3A-Predictive-Modelling-in-Teaching-and-BrooksThompson/2cd4901b07f3562f98e1e56dc5712e8bc03bdc2e> (accessed 04.20.2023).

Acknowledgments: The study was carried out with the support of the TSU Development Program ("Priority-2030"). Authors also express their sincere gratitude to the journal's reviewers for their attention to authors' article.

*The paper was submitted 05.07.2023
Accepted for publication 22.08.2023*
