

Современный мировой ландшафт онлайн-обучения наукам о данных

Научная статья

DOI: 10.31992/0869-3617-2022-31-4-129-147

Никольский Владимир Святославович – д-р филос. наук, проф., гл. редактор журнала «Высшее образование в России», SPIN-code: 7196-8065, ORCID: 0000-0002-4290-1443, v.s.nikolskij@mospolytech.ru

Московский государственный политехнический университет, Москва, Россия

Адрес: 107023, г. Москва, ул. Б. Семёновская, 38

Лукашенко Марианна Анатольевна – д-р экон. наук, проф., зав. кафедрой «Корпоративная культура», SPIN-code: 1393-2243, ORCID: 0000-0001-8707-486X, mlukashenko@synergy.ru

Московский финансово-промышленный университет «Синергия», Москва, Россия

Адрес: 105318, г. Москва, Измайловский вал, 2

Шарова Екатерина Андреевна – канд. экон. наук, зам. руководителя Центра координации исследований, SPIN-code: 8611-5298, ORCID: 0000-0002-7450-7598, ekaterina.sharova@riss.ru

Российский институт стратегических исследований, Москва, Россия

Адрес: 125413, г. Москва, ул. Флотская, 15Б

Аннотация. Науки о данных (data science) показывают высокую динамику в качестве новой отрасли прикладного знания, а также новой области обучения. Кроме того, активно развивается соответствующая сфера образовательных исследований. Вместе с тем большая часть научных публикаций направлена на изучение конкретных вопросов, связанных с содержательным наполнением программ и их методическим обеспечением. Более широкий и особенно международный контекст остаётся без необходимого внимания исследователей.

В этой связи целью нашего исследования явилось обобщение и систематизация сведений о программах обучения в области наук о данных, представленных на онлайн-платформах основных макрорегионов: Америки, Европы и Азии. Для этого мы выяснили, из каких элементов состоит корпус программ обучения наукам о данных, как на образовательных платформах распределены курсы по странам и организациям, предоставляющим курсы (провайдерам), как курсы распределены по уровню образования и продолжительности обучения. На основе полученных данных мы провели сравнительное межрегиональное исследование образовательных программ, представленных на онлайн-платформах.

Полученные данные позволили сделать выводы о специфике мирового ландшафта онлайн-обучения наукам о данных, а также определить специфику российского сегмента и сформулировать рекомендации для решения значимых задач отечественной экономики с помощью онлайн-обучения наукам о данных.

Ключевые слова: науки о данных, онлайн-платформы, MOOK, большие данные, обучение

Для цитирования: Никольский В.С., Лукашенко М.А., Шарова Е.А. Современный мировой ландшафт онлайн-обучения наукам о данных // Высшее образование в России. 2022. Т. 31. №. 4. С. 129-147. DOI: 10.31992/0869-3617-2022-31-4-129-147

The Modern World Landscape of Data Science Online Education

Original article

DOI: 10.31992/0869-3617-2022-31-4-129-147

Vladimir S. Nikolskiy – Dr. Sci. (Philosophy), Prof., Chief Editor of the journal of “Higher Education in Russia”, SPIN-code: 7196-8065, ORCID: 0000-0002-4290-1443, v.s.nikolskij@mospolytech.ru
Moscow State Polytechnic University, Moscow, Russia
Address: 38, B. Semenovskaya str., Moscow, 107023, Russian Federation

Marianna A. Lukashenko – Dr. Sci. (Economics), Prof., Head of the Department of Corporate Culture, SPIN-code: 1393-2243, ORCID: 0000-0001-8707-486X, mlukashenko@synergy.ru
Moscow University for Industry and Finance “Synergy”, Moscow, Russia
Address: 2, Izmaylovskiy val str., Moscow, 105318, Russian Federation

Ekaterina A. Sharova – Cand. Sci. (Economics), Deputy Head of the Research Coordination Center, SPIN-code: 8611-5298, ORCID: 0000-0002-7450-7598, ekaterina.sharova@riss.ru
Russian Institute for Strategic Studies, Moscow, Russia
Address: 15B, Flotskaya str., Moscow, 125413, Russian Federation

Abstract. Data science as an emerging branch of applied knowledge and a new field of study is showing a strong momentum. Besides, the corresponding sphere of educational research is actively developing. At the same time, most of the scientific publications are aimed at studying specific issues related to the content of the programs and their methodological support. The wider context and especially the international perspective are lacking for the necessary attention of researchers.

In this regard, the purpose of our study was to summarize and systematize information about training programs in the field of data science presented on online platforms of the main macro-regions – America, Europe and Asia. For this purpose, we found out what elements the corpus of data science training programs consists of, as well as how courses are distributed on educational platforms by countries, organizational providers, level of education and duration of study. Based on the data obtained, we conducted a comparative interregional study of educational programs presented on online platforms.

The findings made it possible to draw conclusions about the specifics of the global landscape of data science online education, as well as to determine the specifics of the Russian segment and to formulate recommendations for solving significant problems of the domestic economy using data science online education.

Keywords: data science, online platforms, MOOCs, big data, learning

Cite as: Nikolskiy, V.S., Lukashenko, M.A., Sharova, E.A. (2022). The Modern World Landscape of Data Science Online Education. *Vysshee obrazovanie v Rossii = Higher Education in Russia*. Vol. 31, no. 4, pp. 129-147, doi: 10.31992/0869-3617-2022-31-4-129-147 (In Russ., abstract in Eng.).

Введение

Роль данных в экономике, политике и науке исследована достаточно подробно. Большие данные (Big Data) в настоящее время широко используются в ряде отраслей экономики и для принятия политических решений. Большинство исследований касаются применения больших данных, а в последнее время появляются публикации, раскрывающие значение данных и для сферы образования.

В настоящее время известны некоторые исследования, посвящённые специфике обучения работе с большими данными в конкретных областях. Например, рассмотрены вопросы контента программ обучения наукам о данных [1], представлена новая содержательно-методическая линия для модернизации системы профессиональной подготовки будущего экономиста [2], проанализированы аспекты формирования системного мышления у студентов в процессе изучения наук о данных [3]. Вместе с тем тема обучения наукам о данных в широком межрегиональном масштабе изучена не столь глубоко.

Одной из ключевых тем, на наш взгляд, выступает исследование того, кто и как готовит специалистов в области наук о данных. Те, кто попадает сегодня в эту область, обычно переходят в неё из других областей, таких как компьютерные науки, математика, экономика, бизнес, журналистика, психология и другие. «Большинство из тех, кто совершил переход, приобрели свои навыки путём самоподготовки, читая книги и посещая онлайн-курсы... Не через традиционную систему образования»¹. Тем не менее вначале технологические компании, а затем и вузы стали разрабатывать образовательные программы, направленные на обучение наукам о данных, и в настоящее время подготовка специалистов идёт полным ходом как на об-

разовательных платформах, так и в высших учебных заведениях.

Однако подготовка таких специалистов – вызов системе образования, поскольку науки о данных динамично развиваются и трансформируются. Например, одной из составляющих наук о данных являются big data – большие данные, которые выделены в самостоятельный предмет изучения и от которых всё больше компаний ожидают ощутимых результатов в бизнесе². При этом новейшие публикации констатируют переход от больших данных к малым и широким данным и считают его одной из главных тенденций в области данных и аналитики. Поэтому необходимы как регулярная актуализация учебных программ для того, чтобы в процессе обучения студенты могли формировать новые компетенции, так и программы профессиональной переподготовки и дополнительного образования. В высших учебных заведениях должна быть обеспечена готовность к разработке таких программ. В противном случае вузы окажутся неконкурентоспособными по отношению к бизнес-провайдерам, размещающим привлекательные программы на образовательных платформах и не сталкивающимся с тем уровнем бюрократизации, с которым сталкиваются университеты.

В силу того, что в современных условиях удалённое обучение является мейнстримом в образовании, мы сочли наиболее актуальным проанализировать российские и зарубежные образовательные платформы на предмет наличия программ, направленных на обучение наукам о данных, а также выявления основных игроков на рынке образовательных продуктов в данной сфере. В этой связи целью исследования явилось обобщение и систематизация сведений о программах обучения в области наук о данных, представленных на онлайн-платформах ос-

¹ Zaire Cb. Is Data Science Still a Rising Career in 2021 // Medium. 2021. Jan 27. URL: <https://towardsdatascience.com/is-data-science-still-a-rising-career-in-2021-722281f7074c> (дата обращения: 26.03.2022).

² Стельмах С. Как усовершенствовать работу с большими данными и аналитику в 2022 году // itWeek. 2021. 27.12. URL: <https://www.itweek.ru/bigdata/article/detail.php?ID=221761> (дата обращения: 26.03.2022).

новых макрорегионов: Америки, Европы и Азии.

Для этого нам потребовалось выяснить, из каких элементов состоит корпус программ обучения наукам о данных, как на образовательных платформах распределены курсы по странам и организациям, представляющим курсы (провайдерам), как курсы распределены по уровню образования и продолжительности обучения. Наконец, на основе полученных данных было проведено сравнительное межрегиональное исследование образовательных программ, представленных на онлайн-платформах.

Обзор литературы

В этом разделе мы покажем, в каких основных направлениях движутся исследования вопросов обучения в области наук о данных, а также обоснуем актуальность нашего исследования для современных научных дискуссий по этой тематике.

Научные публикации на тему обучения наукам о данных активизировались с появлением в первой половине 2010-х гг. в вузах Северной Америки соответствующих программ и следом за ними – программ по наукам о данных на образовательных онлайн-платформах. В 2014 г. в США насчитывалось 42 программы – от бакалавриата до аспирантуры [1]. Кроме того, большой интерес вызывал и до сих пор вызывает вопрос об управлении в образовании с помощью наук о данных [4].

Прежде всего, необходимо отметить, что в академической среде только начинает складываться консенсус относительно того, как должна выглядеть учебная программа в области наук о данных [5; 6]. Каково должно быть соотношение теоретического и практического обучения, возможно ли индивидуальное обучение или обязательным компонентом должна выступать командная работа? Какими компетенциями должен владеть выпускник программы? [7]. Эти вопросы до сих пор являются дискуссионными. Впрочем, некоторые подходы, апробиро-

ванные в 2014–2015 гг. на образовательных платформах, стали де-факто стандартом онлайн-обучения в области наук о данных [8]. Не устоялись и формы включения наук о данных в образовательные программы. Так, например, науки о данных могут встраиваться в существующие образовательные программы, например, по статистике [9], биотехнологиям [10] или развиваться в качестве самостоятельного направления.

Важной составляющей современных исследований о будущем наук о данных является дискуссия о междисциплинарном и прикладном характере этого направления. В силу того, что в академическом мире науки о данных институционально связаны с кафедрами (департаментами) статистики, практическая направленность выражена не так явно, как на этом настаивают работодатели [11]. Излишняя теоретическая направленность обучения в университетах позволяет компаниям EdTech-сектора предлагать конкурентоспособные решения, связанные с проектным обучением или обучением на основе тематических исследований. В научных дискуссиях заметное место занимают публикации, в которых доказывается принципиальная междисциплинарность, характерная как для молодой науки, так для соответствующих программ обучения и обеспечивающая высокие результаты обучения [12].

Здесь важное значение для обучения заключается в том, что науки о данных рассматриваются как единство данных, технологий и людей [1]. Иными словами, для того чтобы заниматься науками о данных, недостаточно владеть данными и технологиями их обработки, а важнейшее значение для её становления имеют поведенческие науки.

Однако, по оценке Джейсона Уэста, стремительные темпы разработки и внедрения образовательных программ в области наук о данных привели к тому, что университеты смогли предложить лишь те программы, которые соответствовали возможностям университета, а не потребностям обучающихся или работодателей [13]. То есть содержание

программ по наукам о данных критически зависит от уровня подготовки и опыта преподавателей [14].

Факт несоответствия описания программ подготовки тому, что действительно является содержанием и результатами программы, был обнаружен и описан ещё в 2016 г. Можно признать, что с того времени мало что изменилось, хотя для более точных суждений требуется отдельное исследование [15].

Непрерывное и динамичное развитие наук о данных является определённым вызовом для преподавателей, поскольку требует от них существенных временных ресурсов на то, чтобы оставаться на переднем крае в этой области. Вместе с тем преподаватели сообщают о том, что одной из самых больших проблем оказывается преподавание аудитории с разнообразной подготовкой, поскольку адекватных методик преподавания в этих условиях недостаточно [5].

В этой связи интерес вызывают новые формы обучения, в частности командное и проектное обучение. Дело в том, что профессиональная деятельность специалиста в области данных является по преимуществу коллективной, а потому и студенты, вероятно, должны учиться в командном формате [16]. Видимо, по этой причине на онлайн-платформах также стараются использовать этот подход.

Кроме того, эффективное обучение работе с данными зависит от понимания обучающимися того, как эти данные будут использоваться и какую пользу от этого получит общество [17].

Тема обучения наукам о данных остаётся практически неизученной российскими учёными. Текущие исследования касаются прежде всего применения анализа данных в управлении образовательными системами [18], реализации конкретных подходов в обучении в области обработки, хранения и анализа больших данных [19]. Однако развитие самой предметной области, интерес к ней и возможности образования в России соответствуют общемировым тенденциям [20].

Таким образом, массив публикаций, посвящённых образованию в области наук о данных, можно условно разделить на два множества, стремящихся ответить на два принципиальных вопроса: «чему учить?» и «как учить?». В исследованиях вопросов обучения в области наук о данных преобладают тематические исследования. Они весьма полезны как для преподавателей, так и для управленческого состава университетов, но эти работы охватывают только небольшой фрагмент реальности. К сожалению, исследований более широкого контекста образовательных программ наук о данных недостаточно. Это касается как русскоязычных, так и англоязычных публикаций. В этой связи мы предприняли сравнительное межрегиональное исследование образовательных программ, представленных на онлайн-платформах.

Методология

В качестве методологии исследования использовался анализ российской и зарубежной научной литературы, интернет-источников, контента образовательных платформ. Согласно международному опыту для получения данных об онлайн-курсах по обучению наукам о данных в январе 2022 г. был осуществлён поиск источников в российских и международных базах публикаций на основе ключевого слова «data» с указанием темы «data science». Для обработки значительного объёма информации были использованы алгоритмы работы с большими данными, такие как Google Chrome-расширения для извлечения данных с веб-страниц – Instant data scraper и Table Capture.

Были использованы такие базы научного цитирования, как Google Scholar, Scopus, eLibrary, Academia, а также контент образовательных платформ массовых открытых онлайн-курсов – MOOC (Massive Open Online Course – MOOC).

Выборка онлайн-платформ для поиска образовательных продуктов по наукам о данных была построена таким образом, что-

бы охватить основные макрорегионы: Америку, Европу и Азию.

В нашем исследовании мы ориентировались на опубликованный список MOOC-платформ по всему миру в 2022 г.³ В качестве объекта исследования были выбраны наиболее репрезентативные образовательные платформы MOOC из основных макрорегионов:

- Америка: Coursera, edX;
- Европа: FutureLearn (Великобритания), France Université Numérique (FUN) (Франция);
- Азия: XuetangX (Китай), SWAYAM (Индия), Гассо (Япония);
- Россия: «Открытое образование» (openedu.ru).

Среди представленных платформ – ведущие глобальные проекты и официальные национальные образовательные платформы.

Для того чтобы лучше понять отечественную специфику, а также в целях разработки рекомендаций для отечественного рынка в выборку также были добавлены такие ведущие EdTech-платформы, как «Нетология», «Яндекс Практикум», Skillbox, Skillfactory, Geek Brains, Stepik.

Ограничения

Мы получили ограниченный срез данных, касающийся только программ, предоставляемых онлайн-платформами. Наше исследование не включало анализ собственно университетских программ, в рамках которых можно наблюдать большое разнообразие предложений. Современные университеты предлагают программы на всех уровнях послесреднего образования – бакалавриата, магистратуры и аспирантуры. В дальнейших исследованиях было бы интересно получить полную картину обучения наукам о данных. Кроме того, сама выборка онлайн-плат-

форм является ограниченной и не позволяет получить исчерпывающие данные обо всех онлайн-программах по наукам о данных в той или иной стране и в том или ином макрорегионе.

Результаты

Coursera – одна из наиболее крупных на сегодняшний день образовательных платформ, аккумулирующая более 7000 курсов, из которых 1455 – по наукам о данных (по состоянию на 30.01.2022 г.). Предлагаются разнообразные виды образовательных программ:

1) магистерские программы (всего семь продолжительностью в большинстве своём два года, однако заявляется и длительность обучения один–три года). В данной группе Россию представляет НИУ ВШЭ (две программы)⁴. Ещё три программы предложены американскими университетами, а провайдерами двух остальных являются вузы Великобритании и Колумбии;

2) программы на присвоение профессиональных сертификатов по наукам о данных (продолжительность – от полутора до восьми месяцев при занятиях от трёх до 14 часов в неделю). В основном это профессиональные сертификаты IBM и Google;

3) программы на присвоение сертификатов MasterTrack® по наукам о данных. Данные программы представляют собой отдельные модули программ магистерской подготовки длительностью четыре–семь месяцев и могут быть в дальнейшем зачтены в процессе обучения на магистерской программе;

4) программы на присвоение сертификата о специализации по направлению наук о

³ *Shah D., Pickard L., Ma R. Massive List of MOOC Platforms Around The World in 2022 // Class Central. 2022. Jan 11. URL: <https://www.classcentral.com/report/mooc-platforms/> (дата обращения: 26.03.2022).*

⁴ Данные представлены по состоянию на 30.01.2022 г. На момент публикации курсы российских вузов недоступны, и российские университеты не значатся в качестве партнёров Coursera в связи с введёнными против России санкциями. Авторами было принято решение представить анализ в данном виде, для того чтобы лучше понять потенциал российских программ в сфере наук о данных.

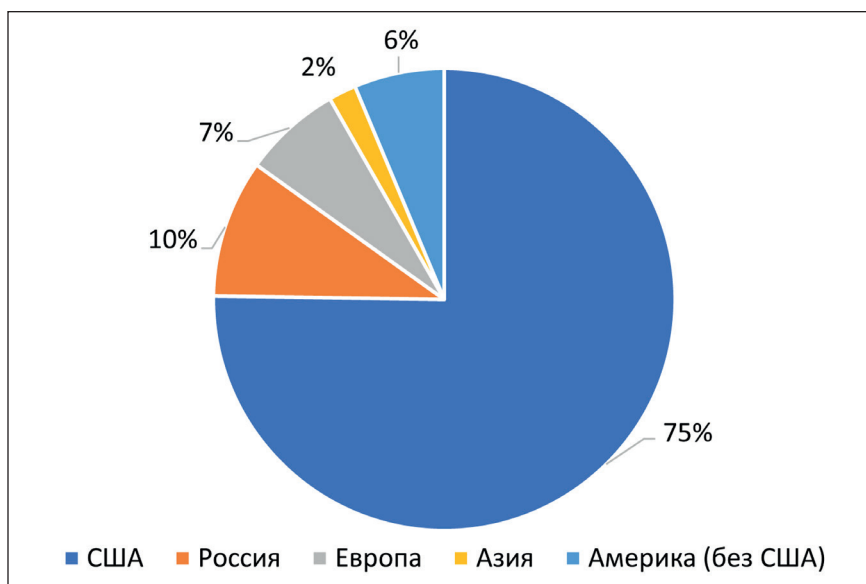


Рис. 1. Распределение курсов по наукам о данных по странам и регионам на образовательной платформе Coursera

Fig. 1. Distribution of data science courses by country and region on the Coursera educational platform

данных. Предполагают прохождение ряда курсов и выполнение проектов;

5) отдельные курсы по направлению наук о данных. В программу курсов включены задания, видеолекции и дискуссионные форумы. По итогам прохождения курса выдаётся электронный сертификат;

6) проект с консультациями (в первоисточнике guided project – управляемый проект) – весьма любопытный образовательный продукт, предлагаемый Coursera. В отличие от продолжительных программ данный продукт предполагает всего двухчасовое взаимодействие с отраслевым экспертом и получение пошаговых инструкций по выполнению проекта. Такие проекты составляют 520 из 1455 образовательных продуктов по наукам о данных и не учитываются нами при проведении аналитики, касающейся распределения курсов по макрорегионам и провайдерам, поскольку нам представляется некорректным сравнивать столь краткосрочный продукт с программами существенно большей длительности. Отметим, однако, что он предоставляется Coursera Project Network,

то есть компаний, осуществляющей обучение силами отраслевых экспертов.

Наибольшее количество курсов по наукам о данных представлено США (702 курса), что составляет 75,08% от общего числа курсов по данному направлению. На втором месте – Россия (90 курсов), что составляет 9,63% (Рис. 1).

Анализ провайдеров выявил картину, подтверждающую нашу гипотезу: несмотря на то, что сами вузы представлены на платформе значительно бóльшим числом, нежели компании, однако по количеству курсов это опережение уже не столь существенное в силу того, что компании предлагают большее число курсов, нежели вузы. Так, если на один вуз в среднем приходится 6,76 курса, то на одну IT-компанию – 18,81, то есть почти в три раза больше. Таким образом, уровень конкуренции вузов и компаний на Coursera весьма высок, во всяком случае, по направлению наук о данных. При этом специализированные IT-компании значительно опережают по числу курсов обучающие компании и прочих игроков (Табл. 1).

Таблица 1

Распределение курсов по наукам о данных по провайдерам на образовательной платформе Coursera

Table 1

Distribution of data science courses by provider on the Coursera educational platform

Провайдер	Количество провайдеров	Количество провайдеров, %	Количество курсов	Количество курсов, %
Вуз	72	68,57	487	52,09
Компания (ИТ)	16	15,24	301	32,19
Компания (обучение)	8	7,62	111	11,87
Бизнес-школа	3	2,86	13	1,39
НИИ	1	0,95	9	0,96
НКО	3	2,86	5	0,53
Институт ЕС (обучение)	1	0,95	5	0,53
Компания (аудит и консалтинг)	1	0,95	4	0,43
Общий итог	105	100,00	935	100,00

Что касается содержания обучения, то большинство онлайн-курсов по наукам о данных включают формирование навыков программирования на языках R, Python, SQL, использование методов анализа данных, статистики и машинного обучения.

EdX – ещё одна крупная американская образовательная платформа, основанная Массачусетским технологическим институтом и Гарвардским университетом в мае 2012 г. 29 июня 2021 г. было объявлено о приобретении edX компанией образовательных технологий 2U⁵. На платформе размещено более трёх тысяч курсов, из них по наукам о данных – 476.

Предлагается обучение по следующим образовательным программам: бакалавриат (только одна программа Лондонского университета); магистратура; курсы; Boot Camps (учебные лагеря); executive education. Курсы по наукам о данных объединены в рубрике «Аналитика данных и статистика». На платформе представлены провайдеры из 26 стран, на первом месте по числу провайдеров и предлагаемых курсов – США. Любопытно,

что страна может быть представлена небольшим количеством провайдеров, но при этом большим числом курсов. Например, Мексике представляют всего два провайдера, однако ими предлагается 16 курсов. Россия представлена Национальным исследовательским университетом ИТМО, Национальным исследовательским ядерным университетом МИФИ и Уральским федеральным университетом (Табл. 2).⁶

Программы Executive education включают рыночно-ориентированные короткие курсы, представленные GetSmarter, брендом 2U, Inc. в сотрудничестве с крупными образовательными учреждениями. Данные курсы включают такие темы, как, например, управленческая аналитика, маркетинговая аналитика, практика применения машинного обучения, монетизация данных, визуализация данных. Наибольшее число курсов представлено Школой менеджмента Слоуна при Массачусетском технологическом институте (США), Лондонской школой экономики (Великобритания) и Университетом Кейптауна (ЮАР).

Анализ провайдеров на платформе EdX выявил картину не столь жёсткой конку-

⁵ *Herpich, N.* edX acquired by education technology company 2U; proceeds to be invested in non-profit // The Harvard Gazette. 2021. June 29. URL: <https://news.harvard.edu/gazette/story/2021/06/edx-acquired-by-education-technology-company-2u/> (дата обращения: 26.03.2022).

⁶ Данные представлены по состоянию на 30.01.2022 г. На момент публикации курсы российских вузов недоступны на EdX в связи с введёнными против России санкциями.

Таблица 2
Распределение курсов по наукам о данных по странам на образовательной платформе EdX
Table 2

Distribution of data science courses by country on the EdX educational platform

Страна	Количество провайдеров	Количество провайдеров, %	Количество курсов	Количество курсов, %
США	38	41,76	259	54,41
Америка (без США)	12	13,19	73	15,33
Европа	17	18,69	65	13,65
Азия	10	11,00	32	6,72
Австралия и Океания	5	5,50	24	5,04
Россия	3	3,30	8	1,68
Международные организации	3	3,30	8	1,68
Африка и Ближний Восток	3	3,30	7	1,47
Общий итог	91	100	476	100

Таблица 3
Распределение курсов по наукам о данных по провайдерам на образовательной платформе EdX
Table 3

Distribution of data science courses by providers on the EdX educational platform

Провайдеры	Количество провайдеров	Количество провайдеров, %	Количество курсов	Количество курсов, %
Вуз	72	79,12	350	73,53
Компания (ИТ)	5	5,49	76	15,97
Компания (промышленность, медицина)	3	3,30	13	2,73
Международная организация	3	3,30	12	2,52
Школа бизнеса	3	3,30	11	2,31
Компания (EdTech)	3	3,30	10	2,10
Международная организация (инициатива ООН)	1	1,10	3	0,63
Государственный орган (департамент образования)	1	1,10	1	0,21
Общий итог	91	100,00	476	100,00

ренции между вузами и компаниями, как на Coursera. Вузы представлены на платформе в наибольшем процентном отношении по сравнению с другими провайдерами (79,12%). Тем не менее, хотя ИТ-компании составляют всего 5,49%, на предлагаемые ими курсы приходится 15,97%, а на одного провайдера – 15,2 курса, в отличие от вузов, где на одного провайдера приходится 4,9 курса (Табл. 3).

FutureLearn – британская платформа цифрового образования, основанная в дека-

бре 2012 г. Компания находится в совместном владении The Open University и SEEK Ltd.⁷ На ней размещены отдельные курсы, серии курсов ExpertTrack и программы на получение степени, всего более 1400 курсов, из них по наукам о данных – 75.

⁷ *Lederman D.* MOOC Platforms' New Model Draws Big Bet from Investors // Inside Higher Ed. 2019. May 22. URL: <https://www.insidehighered.com/digital-learning/article/2019/05/22/investors-bet-big-companies-formerly-known-mooc-providers> (дата обращения: 26.03.2022).

Таблица 4

Распределение курсов по наукам о данных по странам на образовательной платформе FutureLearn

Table 4

Distribution of data science courses by country on the FutureLearn educational platform

Страна	Количество курсов	Количество курсов, %
Филиппины	31	41,33
Великобритания	19	25,33
Европа (кроме Великобритании)	5	6,67
Австралия и Новая Зеландия	5	6,67
США	4	5,33
Коллаборации (Европа, США, Филиппины)	11	14,67
Общий итог	75	100,00

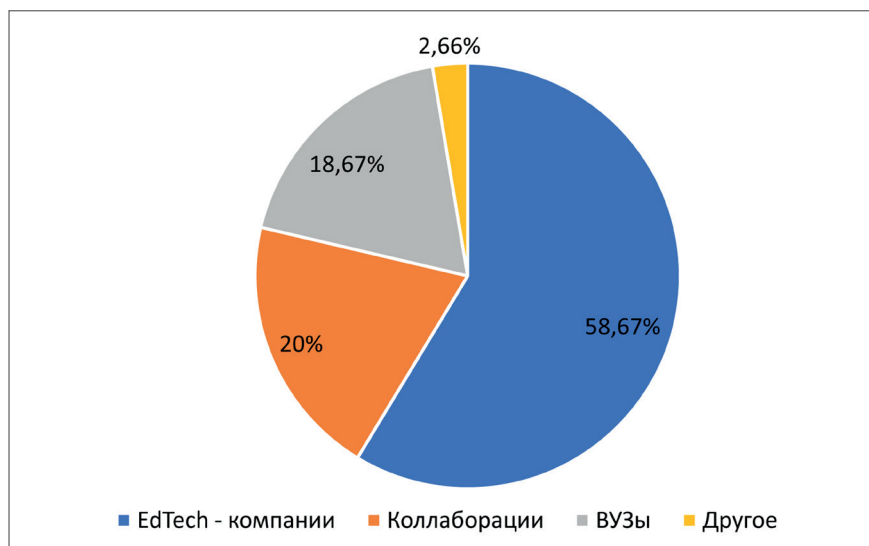


Рис. 2. Распределение курсов по наукам о данных по провайдерам на образовательной платформе FutureLearn

Fig. 2. Distribution of data science courses by providers on the FutureLearn educational platform

ExpertTrack – достаточно новая разработка (2021 г.), представляющая собой подобранную серию курсов по определённой тематике и направленную на более глубокое изучение предметной области. Предлагается три уровня обучения: новичок, средний и продвинутый.

Курсы по наукам о данных предлагают провайдеры из 14 стран, причём первое место по числу курсов занимают Филиппины, что связано с местоположением штаб-квартиры компании CloudSwyft Global Systems, Inc., которая лидирует по числу курсов на плат-

форме (Табл. 4). Отметим, что 11 курсов на платформе представлены коллаборацией провайдеров из разных стран (Европа, США, Филиппины).

Анализируя провайдеров курсов, мы столкнулись с весьма интересным опытом коллаборации провайдеров. Это коллаборации вузов и ИТ-компаний, вузов и органов государственной власти, компаний, занимающихся образовательными технологиями (EdTech), и ИТ-компаний и т.д. Из 10 провайдеров шесть представляют собой коллаборации. При этом наибольшее количество

курсов представлено EdTech-компаниями, но не вузом (Рис. 2).

France Université Numérique (FUN). Цифровой университет Франции (FUN) – французская платформа цифрового образования, на которой представлено порядка 1000 курсов. Более 600 из них архивированы, а около 400 идут в настоящее время либо запланированы на ближайшие месяцы. Предлагается всего одна магистерская программа, и это программа на двух языках по наукам о данных – «Online Master of Science Big Data Intelligence for Human Augmented Reality (eVINAR)». Она открыта для студентов, получивших годовую степень магистра (Master 1) или её эквивалент. Контент программы включает академические онлайн-курсы, курсы, предлагаемые Oracle University, а также дополнительные курсы от Google.

Помимо этого, FUN участвует в Европейском консорциуме MOOK (European MOOC Consortium), в который также входят рассмотренный нами выше британский проект FutureLearn, MiriadaX (Испания и Иbero-Америка), EduOpen (Италия), iMooX и OpenupEd (Европейская ассоциация университетов дистанционного обучения – EAD-TU). В рамках этого проекта FUN в сотрудничестве с ведущими вузами Франции разработал несколько программ дополнительного образования GRADEO. Сертификация по программам GRADEO признаётся во всех странах, участвующих в Европейском консорциуме MOOK.

FUN предлагает две программы GRADEO по наукам о данных:

- расширенное программирование sql;
- большие данные и искусственный интеллект.

В программу GRADEO «Большие данные и искусственный интеллект» входят следующие курсы:

- 1) академические курсы от Institute of Advanced Industrial Technologies (ESTIA):
 - искусственный интеллект MLDL;
 - распределённое управление большими данными;

– реляционное и объектное программирование SQL;

2) профессиональный курс ORACLE:

- машинное обучение в ORACLE CLOUD.

Помимо программ GRADEO, на платформе предлагаются отдельные курсы по наукам о данных. Провайдерами курсов являются высшие учебные заведения Франции, и обучение идёт на французском языке.

SWAYAM – национальная платформа MOOK в Индии. На ней предлагается более 2000 курсов, которые преподаются около 1300 инструкторами из более чем 135 индийских университетов.

Обучение в SWAYAM состоит из четырёх частей/квадрантов: e-Tutorial, электронный контент, дискуссионный форум, ассессмент [21]. Первый квадрант – это непосредственно обучение: видеолекции, презентации и т.д. Второй квадрант представляет собой электронные книги, тематические исследования и т.д. Третий квадрант – площадка для общения учащихся и преподавателя, ответы на вопросы. Четвёртый – тесты для самооценки плюс часто задаваемые вопросы и ответы на них. Как мы видим, структура достаточно классическая для электронных курсов.

SWAYAM – это государственная программа Индии. Для управления контентом курса Министерством развития людских ресурсов (MHRD) назначены девять национальных координаторов, каждый из которых отвечает за свою область.

По ключевому слову «data» на платформе выдаётся 25 курсов продолжительностью 4–16 недель. Наибольшее количество программ посвящено таким темам, как анализ данных, системы управления базами данных, сбор данных, программирование с использованием языков R и Python.

Gacco – это веб-сервис на японском языке, с помощью которого любой желающий может бесплатно прослушать курсы престижных японских университетов. Платформа создана по инициативе NTT DOCOMO и NTT Knowledge Square (NTT – крупный телекоммуникационный провайдер Японии) и Со-

вета по продвижению массового открытого онлайн-образования Японии (Japan Massive Open Online Education Promotion Council, JMOOC). В состав Совета входят университеты, исследовательские центры и ассоциации в области компьютерных наук и образования, а также представители бизнеса (несколько компаний в сфере разработки ПО, ведущий оператор мобильной связи, венчурная компания, ряд компаний в сфере обучения).

В настоящее время возможна регистрация на 70 курсов, из них по наукам о данных и статистике мы выявили 10 курсов продолжительностью от одной до 11 недель. Их провайдерами являются японские университеты (6 из 10), компании (Google и японская Data Vehicle Co., Ltd, занимающаяся консалтингом, обучением и информационными услугами в сфере анализа данных), а также НИИ (Институт статистической математики) и государственное учреждение (Бюро статистики Министерства внутренних дел и коммуникаций). Большинство курсов посвящено статистическим методам анализа данных. В архиве с 2014 г. 412 курсов, в том числе по наукам о данных и статистике – около 80.

XuetangX – крупная китайская платформа МООК, функционирующая во взаимодействии с Исследовательским центром онлайн-образования Министерства образования КНР.

В настоящее время XuetangX предоставляет доступ к более чем 3000 курсов от ведущих университетов страны и зарубежья, таких как Университет Цинхуа, Пекинский университет, Университет Фудань, Китайский университет науки и технологий, Стэнфордский университет и Калифорнийский университет в Беркли.

В результате поиска на английском языке нами выявлено на платформе 22 курса по наукам о данных. Провайдерами подавляющего большинства курсов (18) являются китайские вузы. Кроме того, на платформе представлено два курса от университета из Южной Кореи, один курс от IT-компании (Microsoft) и один курс от коллаборации ки-

тайского вуза и международной организации (мастер-класс по использованию искусственного интеллекта для целей устойчивого развития Университета Цинхуа совместно с Программой развития ООН). Наиболее часто предлагаются курсы по таким темам, как структуры данных и разработка алгоритмов, статистика, математические основы наук о данных (матричный анализ, комбинаторика), большие данные и городское планирование.

«Открытое образование» (*openedu.ru*) – российская образовательная платформа, созданная Ассоциацией «Национальная платформа открытого образования» и предлагающая онлайн-курсы по базовым дисциплинам, изучаемым в российских вузах. В настоящее время на платформе содержится 843 курса по разным направлениям подготовки.

Курсы доступны бесплатно, однако при желании зачесть какой-либо пройденный курс при освоении образовательной программы бакалавриата или специалитета в вузе необходимо получить сертификат. Это возможно при условии успешного прохождения контрольных мероприятий онлайн-курса и является платной услугой. Данная практика распространена почти на всех образовательных платформах и не является российской особенностью.

На платформе девять провайдеров предлагают 40 курсов, соответствующих тематике наук о данных. Кроме того, представлена программа дополнительного профессионального образования «Аналитика данных», разработанная университетом ИТМО и рассчитанная на 19 недель. Программа включает такие курсы, как первичная обработка данных, хранение больших данных, статистический анализ данных, элементы теории вероятностей (факультативно).

Провайдерами курсов являются исключительно российские университеты (Рис. 3).

Помимо *openedu.ru*, на российском рынке онлайн-образования функционирует множество EdTech-компаний, и предложение

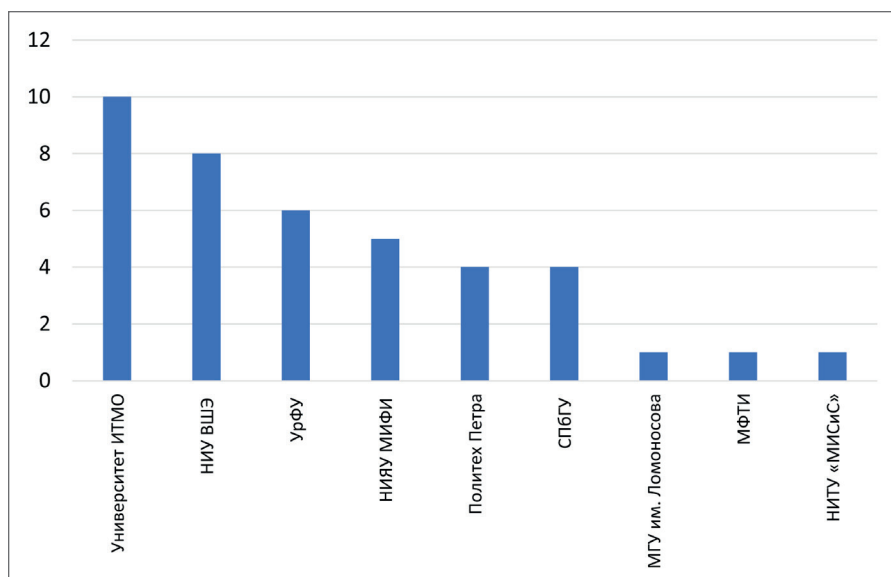


Рис. 3. Число курсов по наукам о данных по провайдерам на платформе «Открытое образование»

Fig. 3. Number of data science courses by provider on the Open Education platform

образовательных продуктов весьма объёмно и разнообразно. Рассмотрим некоторые из них.

«Нетология» – образовательная платформа, созданная для обучения взрослых цифровым профессиям. Всего представлено более 350 продуктов, в том числе 47 – в области аналитики данных, из них 14 так называемых профессий (программ длительностью 9–20 месяцев), две специализации (5–11 месяцев), 25 курсов (10 дней – 6 месяцев) и шесть ознакомительных занятий.

Некоторые программы созданы совместно с технологичными компаниями. Например, профессия «Data Scientist» сформирована в партнёрстве с компанией Dodo Brands и включает решение реального кейса от этой компании в рамках дипломного проекта.

К особенностям реализуемых программ относятся кейсы в портфолио (декларируется более 16), доступ в профессиональные сообщества, участие в конкурсах Kaggle (соревновательный анализ данных на международной конкурсной площадке, на выходе получение рейтинга, играющего существенную роль при устройстве на работу), содействие

в трудоустройстве, помощь в адаптации на новом месте работы.

«Яндекс Практикум» – сервис онлайн-образования от компании «Яндекс». Всего к настоящему времени на платформе размещено около 50 курсов, из них по наукам о данных – две программы повышения квалификации («Аналитик данных» и «Специалист по data science», рассчитанные на шесть и восемь месяцев соответственно) и две программы профессиональной переподготовки («Аналитик данных плюс» и «Специалист по data science плюс», рассчитанные на 12 и 16 месяцев соответственно).

Немаловажной особенностью программы является возвращение денег за оставшуюся часть программы в любой момент её прохождения по запросу без объяснения причины, а также возможность ухода в академический отпуск.

Skillbox – образовательная платформа, объединяющая ведущих экспертов и практиков рынка, методистов и продюсеров образовательного контента. Позиционируется как обучающая профессиям с нуля. Всего на платформе зарегистрирован 521 курс, из них

в сфере анализа данных – пять профессий, рассчитанных на 15–24 месяца, и 12 курсов длительностью 3–12 месяцев.

Совместно с РАНХиГС разработана программа онлайн-бакалавриата «Data Science & Machine Learning» по направлению 09.03.03 Прикладная информатика, профиль «Анализ данных». Особенностью программы является возможность получения второго диплома Университета Гренобль Альпы (Франция).

SkillFactory – школа онлайн-профессий. Всего на платформе более 60 курсов. Образовательные программы по наукам о данных и аналитике данных представлены 16 продуктами, из которых три программы длительностью освоения 24 месяца, два полных курса длительностью от 6 до 12 месяцев и 11 курсов длительностью от двух до семи месяцев.

Кроме того, предлагаются две онлайн-магистратуры длительностью 24 месяца:

- «Наука о данных» совместно с МИСиС, партнёрами программы также являются Московское агентство инноваций, компания DANONE, французская облачная платформа Девхаб, компания First Line Software;
- «Инженерия машинного обучения» совместно с Уральским федеральным университетом.

GeekBrains – российская компания в сфере онлайн-образования в составе Mail.ru Group. В сфере аналитики данных на данной платформе представлено 18 программ длительностью 6–36 месяцев. Особенности обучения по любой программе являются:

- гарантия трудоустройства через девять месяцев по IT-профессиям;
- диплом о профессиональной переподготовке;
- возможность создания бизнеса;
- бессрочный доступ к программе;
- доступ к дополнительным курсам GeekBrains и Skillbox.

Stepik – российская образовательная платформа и конструктор бесплатных и платных открытых онлайн-курсов и уроков. Даёт возможность каждому зарегистрированному пользователю создавать интерак-

тивные обучающие уроки и онлайн-курсы, используя видео, тексты и задачи с автоматической проверкой и моментальной обратной связью. Кроме того, представляет собой площадку для проведения олимпиад. Целевыми аудиториями являются школьники (в основном курсы по подготовке к ЕГЭ), студенты, начинающие специалисты.

Авторы онлайн-курсов на платформе – физические и юридические лица. Среди организаций, выпустивших онлайн-курсы на платформе, – компании «Яндекс», Samsung, Mail.ru и целый ряд вузов. В подборке по теме «Статистика и анализ данных» мы обнаружили 40 курсов продолжительностью обучения от 2 до 45 часов.

Обсуждение

Рассмотрим более подробно региональную специфику. Она выражается скорее не в контенте представленных образовательных продуктов, а в их многообразии в зависимости от степени развития онлайн-платформы. Так, наиболее мощными на сегодня являются американские платформы Coursera и EdX. На них в совокупности представлено девять видов образовательных продуктов по наукам о данных – от двухчасового проекта с консультациями до двухгодичной программы магистерской подготовки и Executive Education. Отдельного внимания заслуживают программы Boot Camps (учебные лагеря). Это интенсивные иммерсивные («с погружением») программы, сочетающие в себе рыночно-ориентированные учебные планы, обучение на основе проектов и глобальные сообщества обучающихся. В результате прохождения такой программы у обучающихся формируется портфолио работ, что весьма полезно для трудоустройства.

Важную роль в обучении наукам о данных на американских образовательных платформах играют провайдеры из IT-компаний, чьё присутствие формирует весьма выраженную конкурентную среду и обуславливает необходимость практико-ориентированного подхода к обучению со стороны учебных заведений.

Европа в нашем исследовании представлена ещё одним заслуженным лидером среди образовательных платформ – британской FutureLearn. Наиболее интересным в организации обучения на платформе FutureLearn, на наш взгляд, является интенсивная коллаборация провайдеров, а также то обстоятельство, что наибольшее количество курсов представлено EdTech-компанией, а не университетами. Мы считаем, что коллаборации вузов и IT-компаний в сфере обучения наукам о данных являются наиболее ориентированной на результат, продуктивной и, следовательно, перспективной формой оказания образовательных услуг, способствующей большей практической направленности обучения, применению проектного подхода и трудоустройству успешных студентов.

Что касается рассмотренной выше франкоязычной платформы FUN, то особенностью в данном случае является её участие в Европейском консорциуме MOOC, занимающегося разработкой совместных программ дополнительного образования и последующим признанием сертификации по ним. Например, сертификация по программам Gradeo, реализуемым FUN, признаётся во всех странах, участвующих в Европейском консорциуме MOOC.

Азию в нашем исследовании представляют платформы SWAYAM (государственная программа Индии), Gacco (Япония) и XuetangX (Китай).

Характерно, что в Индии для управления контентом онлайн-курсов Министерством развития людских ресурсов назначаются национальные координаторы, например Всеиндийский совет по техническому образованию (AICTE), Консорциум по образовательным коммуникациям (SEC) и Национальный открытый университет имени Индиры Ганди (IGNOU). Ключевыми провайдерами курсов по наукам о данных являются Индийские технологические институты (ИТИ) – автономные государственные технические и исследовательские институты и индийские университеты.

Особенностью японского обучения наукам о данных является то, что все курсы на платформе представлены исключительно на японском языке и, следовательно, ориентированы на внутреннего пользователя. Провайдерами курсов в большинстве своём являются японские университеты, однако присутствуют и коллаборация университета с Google и другие организации.

Китайская образовательная платформа, наоборот, нацелена на развитие в качестве международной. Наличие множества курсов на английском языке свидетельствуют о стремлении Китая со временем войти в число мировых лидеров в сфере онлайн-образования. Однако подавляющее большинство провайдеров программ по наукам о данных – китайские университеты (за редким исключением).

Россия в нашем исследовании представлена образовательной платформой «Открытое образование» (openedu.ru), созданной Ассоциацией «Национальная платформа открытого образования», а также платформами «Нетология», «Яндекс Практикум», Skillbox, Skillfactory, Geek Brains, Stepik.

В сравнении с зарубежными платформами в отношении обучения наукам о данных Openedu.ru выглядит весьма достойно. К особенностям платформы следует отнести то, что все представленные курсы разработаны в соответствии с требованиями федеральных государственных образовательных стандартов и релевантны требованиям к результатам освоения образовательных программ, реализуемых в вузах.

Что касается EdTech-компаний, то их безусловным достоинством является практическая направленность обучения, широкий выбор образовательных продуктов – от коротких курсов, направленных на отработку новых инструментов, до освоения новой специальности с нуля. Важную роль играет возможность формирования портфолио на основании работы с реальными кейсами и в особенности опция трудоустройства, предла-

гаемая некоторыми платформами. Находит своё развитие также коллаборация бизнеса и университетов. На наш взгляд, очень интересный опыт демонстрирует платформа Stepiк, на которой возможно создание востребованных онлайн-курсов не только компаниями и вузами, но и частными экспертами.

Выводы и заключение

1. Мировой ландшафт онлайн-обучения наукам о данных имеет ярко выраженную национальную принадлежность. Несмотря на попытку США создать глобальные платформы, ландшафт обучения в области наук о данных представлен прежде всего национальными системами. Глобальность инициатив США выглядит довольно однобоко, поскольку практически половина всех партнёров Coursera принадлежит США, а число их курсов по наукам о данных превышает 75%. К тому же Coursera напрямую зависит от американского законодательства: в соответствии с правилами экспортного контроля США компания не может предлагать услуги пользователям в регионах, против которых введены американские санкции. Так, для Ирана, Крыма, Кубы, Северной Кореи, Сирии и Судана закрыт доступ к ряду курсов (в сфере углублённого изучения науки, технологий, инженерии и математики), пользователи из указанных стран могут также столкнуться с блокировкой при оплате открытых курсов⁸. Начиная с 4 марта 2022 г. оплата обучения недоступна для всех пользователей из России, при этом российские курсы исключены из актуального каталога⁹.

2. Как показало наше исследование, практически на всех образовательных плат-

формах независимо от макрорегиона представлено большое число образовательных продуктов по наукам о данных по схожей тематике. Россия находится в высококонкурентной среде, и для реализации стратегии привлечения IT-специалистов для работы в нашей стране необходимо развивать экспорт образования, создавая курсы в том числе на иностранных языках и размещая их на отечественных и зарубежных образовательных платформах, в частности на китайской. Россия испытывает острую потребность в специалистах по работе с данными, а различные факторы, в том числе демографические, не позволяют надеяться на быструю подготовку необходимого для экономики числа отечественных специалистов. Необходим приток IT-специалистов из-за рубежа. Зарубежные платформы, в тех случаях, когда это возможно, следует использовать для реализации стратегии привлечения рабочей силы для работы в нашей стране, создавая понятные траектории от обучения к трудоустройству. В этих же целях представляется целесообразным, по опыту Китая, развивать международную составляющую отечественных образовательных платформ, привлекая к участию в них университеты, компании и отдельных специалистов из государств – стратегических партнёров России, главным образом – стран-участниц ЕАЭС и ШОС.

3. В настоящее время на российских онлайн-платформах представлен широкий выбор образовательных продуктов в области наук о данных (в общей сложности 135 программ различного уровня сложности), ориентированных на решение разных задач – от формирования и развития конкретного навыка до овладения новой профессией, приобретения степени бакалавра/магистра, профессиональной переподготовки или повышения квалификации. Круг провайдеров, осуществляющих обучение наукам о данных, весьма широк, хотя на сегодняшний день ключевыми провайдерами образовательных услуг все же являют-

⁸ International restrictions // Coursera. 2021. Sep. 17. URL: https://www.coursera.support/s/article/208280116-International-restrictions?language=en_US (дата обращения: 26.03.2022).

⁹ Coursera's response to the humanitarian crisis in Ukraine // Coursera. 2022. March 4. URL: <https://blog.coursera.org/coursera-response-to-the-humanitarian-crisis-in-ukraine/> (дата обращения: 26.03.2022).

ся вузы. Вместе с тем для России важной особенностью является активная работа EdTech-компаний, предлагающих весьма привлекательный набор программ для обучения наукам о данных. В этой связи для того, чтобы решать как свои, так и государственные задачи, вузам следует в первую очередь конкурировать не между собой и не с коммерческими провайдерами, а за студентов, в том числе иностранных, создавая успешные коллаборации с работодателями и EdTech-компаниями.

Пути дальнейших исследований мы видим прежде всего в изучении наук о данных как области образовательных исследований. Как мы писали в обзоре литературы, исследования делятся на два больших лагеря и стремятся ответить на два важнейших вопроса: «чему учить» и «как учить». В этой связи нас интересуют два направления будущих исследований: 1) специфические методики обучения междисциплинарным областям знаний и деятельности и 2) включение этической тематики в образовательные программы в области наук о данных.

Литература

1. Song I.-Y., Zhu Y. Big Data and Data Science: What Should We Teach? // *Expert Systems*, 2016. Vol. 33. Issue 4. P. 364–373. DOI: 10.1111/exsy.12130
2. Власов Д.А., Карасев П.А., Синчуков А.В. Возможности новой содержательно-методической линии «Анализ больших данных» для модернизации системы профессиональной подготовки будущего экономиста // *Статистика и Экономика*. 2021. Т. 18. № 5. С. 60–70. DOI: <https://doi.org/10.21686/2500-3925-2021-5-60-70>
3. Пыркина О.Е. Методология формирования системного мышления у студентов в процессе изучения курсов data science // *Современная математика и концепции инновационного математического образования*. 2020. Т. 7. № 1. С. 442–446. URL: <https://elibrary.ru/item.asp?id=43009266> (дата обращения: 26.03.2022).
4. Williamson B. Governing methods: Policy innovation labs, design and data science in the digital governance of education // *Journal of Educational Administration and History*. 2015. Vol. 47. Issue 3. P. 251–271. DOI: 10.1080/00220620.2015.1038693
5. Schwab-McCoy A., Baker C., Gasper R. Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years // *Journal of Statistics and Data Science Education*. 2021. Vol. 29. Issue sup1: Computing in the Statistics and Data Science Curriculum. P. S40–S50. DOI: 10.1080/10691898.2020.1851159
6. Irizarry R.A. The role of academia in data science education // *Harvard Data Science Review*. 2020. Issue 2.1. DOI: 10.1162/99608f92.dd363929
7. Demchenko Y., Belloum A., Wiktorowski T. Edison Data Science Framework: Part 1. Data Science Competence Framework (CF-DS), Release 1 // *Zenodo*. 2016. Oct 10. DOI: 10.5281/zenodo.167585
8. Kross S., Peng R.D., Caffo B.S., Gooding I., Leek J.T. The Democratization of Data Science Education // *American Statistician*. 2020. Vol. 74. Issue 1. P. 1–7. DOI: 10.1080/00031305.2019.1668849
9. Hardin J., Hoerl R., Horton N.J., Nolan D., Baumer B., Hall-Holt O., Murrell P., Peng R., Roback P., Temple Lang D., Ward M.D. Data Science in Statistics Curricula: Preparing Students to Think with Data // *The American Statistician*. 2015. Vol. 69. Issue 4. DOI: 10.1080/00031305.2015.1077729
10. Emery N.C., Crispo E., Supp S.R., Farrell K.J., Kerkhoff A.J., Bledsoe E.K., O'Donnell K.L., McCall A.C., Aiello-Lammens M.E. Data Science in Undergraduate Life Science Education: A Need for Instructor Skills Training // *BioScience*. 2021. Vol. 71. Issue 12. P. 1274–1287. DOI: 10.1093/biosci/biab107
11. Hicks S.C., Irizarry R.A. A Guide to Teaching Data Science // *American Statistician*. 2018. Vol. 72. Issue 4. P. 382–391. DOI: 10.1080/00031305.2017.1356747
12. Asamoab B., Doran B., Schiller S. (2020) Interdisciplinarity in Data Science Pedagogy: A Foundational Design // *Journal of Computer Information Systems*. 2020. Vol. 60. Issue 4. P. 370–377. DOI: 10.1080/08874417.2018.1496803
13. West J. Teaching data science: An objective approach to curriculum validation // *Computer Science Education*. 2018. Vol. 28. Issue 2. P. 136–157. DOI: 10.1080/08993408.2018.1486120
14. Yan D., Davis G.) A First Course in Data Science // *Journal of Statistics Education*. 2019. Vol. 27. Issue 2. P. 99–109. DOI: 10.1080/10691898.2019.1623136

15. Tang R., Sae-Lim W. Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus // *Education for Information*. 2016. Vol. 32. No. 3. P. 269–290. DOI: 10.3233/EFI-160977
16. Vance E. Using Team-Based Learning to Teach Data Science // *Journal of Statistics and Data Science Education*. 2021. Vol. 29. Issue 3. P. 277–296. DOI: 10.1080/26939169.2021.1971587
17. Дрябин А.А., Попов А.А. Обучение дата-грамотности: содержание и конкурентные преимущества образовательной программы // *Информационное общество*. 2021. № 1. С. 21–29. DOI: 10.52605/16059921_2021_01_21
18. Захарова И.Г. Big Data и управление образовательным процессом // *Вестник Тюменского государственного университета. Гуманитарные исследования*. 2017. Т. 3. № 1. С. 210–219. DOI: 10.21684/2411-197X-2017-3-1-210-219
19. Белов М.А., Гришко С.И., Лишилин М.В., Оситов П.А., Черемисина Е.Н. Стратегия подготовки ИТ-специалистов с применением инновационного учебного дата-центра “Виртуальная компьютерная лаборатория” для эффективного решения задач цифровой трансформации и акселерации цифровой экономики // *Современные информационные технологии и ИТ-образование*. 2021. Т. 17. № 1. С. 134–144. DOI: 10.25559/SITITO.17.202101.703
20. Наумов В.Н. Анализ результатов опроса Kaggle о состоянии науки о данных и машинном обучении в России // *Научные труды Северо-Западного института управления РАНХиГС*. 2020. Т. 11. № 4 (46). С. 172–184. URL: https://spb.ranepa.ru/wp-content/uploads/2021/03/nt_11_4.pdf (дата обращения: 26.03.2022).
21. Anuva S. Analytical Study of SWAYAM // *IJRAR – International Journal of Research and Analytical Reviews*. 2018. Vol. 5. Issue 3. URL: http://ijrar.com/upload_issue/ijrar_issue_1471.pdf (дата обращения: 26.03.2022).

Статья поступила в редакцию 27.02.22

Принята к публикации 26.03.22

References

1. Song, I.-Y., Zhu, Y. (2016). Big Data and Data Science: What Should We Teach? *Expert Systems*. Vol. 33, issue 4, pp. 364–373, doi: 10.1111/exsy.12130
2. Vlasov, D.A., Karasev, P.A., Sinchukov, A.V. (2021). Opportunities for a New Content and Methodological Line “Big Data Analysis” to Modernize the Training System of the Future Economist. *Statistika i ekonomika = Statistics and Economics*. Vol. 18, no. 5, pp. 60–70, doi: <https://doi.org/10.21686/2500-3925-2021-5-60-70> (In Russ., abstract in Eng.).
3. Pyrkina, O.E. (2020). Methodology for Operational Thinking Development for the Students Studying Data Science. *Sovremennaya matematika i kontseptsii innovatsionnogo matematicheskogo obrazovaniya* [Modern Mathematics and Concepts of Innovative Mathematics Education]. Vol. 7, no. 1, pp. 442–446. Available at: <https://elibrary.ru/item.asp?id=43009266> (accessed 26.03.2022). (In Russ., abstract in Eng.).
4. Williamson, B. (2015). Governing Methods: Policy Innovation Labs, Design and Data Science in the Digital Governance of Education. *Journal of Educational Administration and History*. Vol. 47, issue 3, pp. 251–271, doi: 10.1080/00220620.2015.1038693
5. Schwab-McCoy, A., Baker, C., Gasper, R. (2021). Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years. *Journal of Statistics and Data Science Education*. Vol. 29, issue sup1: Computing in the Statistics and Data Science Curriculum, pp. S40–S50, doi: 10.1080/10691898.2020.1851159
6. Irizarry, R.A. (2020). The Role of Academia in Data Science Education. *Harvard Data Science Review*. Vol. 2, issue 1, doi: 10.1162/99608f92.dd363929
7. Demchenko, Y., Belloum, A., Wiktorski, T. (2016). Edison Data Science Framework-Work: Part 1. Data Science Competence Framework (CF-DS), Release 1. *Zenodo*. Oct 10, doi: 10.5281/zenodo.167585

8. Kross, S., Peng, R.D., Caffo, B.S., Gooding, I., Leek, J.T. (2020). The Democratization of Data Science Education. *American Statistician*. Vol. 74, issue 1, pp. 1-7, doi: 10.1080/00031305.2019.1668849
9. Hardin, J., Hoerl, R., Horton, N.J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, P., Temple Lang, D., Ward, M.D. (2015). Data Science in Statistics Curricula: Preparing Students to Think with Data. *The American Statistician*. Vol. 69, issue 4, doi: 10.1080/00031305.2015.1077729
10. Emery, N.C., Crispo, E., Supp, S.R., Farrell, K.J., Kerkhoff, A.J., Bledsoe, E.K., O'Donnell, K.L., McCall, A.C., Aiello-Lammens, M.E. (2021). Data Science in Undergraduate Life Science Education: A Need for Instructor Skills Training. *BioScience*. Vol. 71, issue 12, pp. 1274-1287, doi: 10.1093/biosci/biab107
11. Hicks, S.C., Irizarry, R.A. (2018). A Guide to Teaching Data Science. *American Statistician*. Vol. 72, issue 4, pp. 382-391, doi: 10.1080/00031305.2017.1356747
12. Asamoah, B., Doran, B., Schiller, S. (2020). Interdisciplinarity in Data Science Pedagogy: A Foundational Design. *Journal of Computer Information Systems*. Vol. 60, issue 4, pp. 370-377, doi: 10.1080/08874417.2018.1496803
13. West, J. (2018). Teaching Data Science: An Objective Approach to Curriculum Validation. *Computer Science Education*. Vol. 28, issue 2, pp. 136-157, doi: 10.1080/08993408.2018.1486120
14. Yan, D., Davis, G. (2019). A First Course in Data Science. *Journal of Statistics Education*. Vol. 27, issue 2, pp. 99-109, doi: 10.1080/10691898.2019.1623136
15. Tang, R., Sae-Lim, W. (2016). Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus. *Education for Information*. Vol. 32, no. 3, pp. 269-290, doi: 10.3233/EFI-160977
16. Vance, E. (2021). Using Team-Based Learning to Teach Data Science. *Journal of Statistics and Data Science Education*. Vol. 29, issue 3, pp. 277-296, doi: 10.1080/26939169.2021.1971587
17. Deryabin, A.A., Popov, A.A. (2021). Teaching Data Literacy: Curriculum Content and Competitive Advantage. *Informatsionnoe obschestvo = Information Society*. No. 1, pp. 21-29, doi: 10.52605/16059921_2021_01_21 (In Russ., abstract in Eng.).
18. Zakharova, I.G. (2017). Big Data and Educational Process Management. *Vestnik Tyumenskogo gosudarstvennogo universiteta. Gumanitarnye issledovaniya. Humanitates = Tyumen State University Herald. Humanities Research. Humanitates*. Vol. 3, no. 1, pp. 210-219, doi: 10.21684/2411-197X-2017-3-1-210-219 (In Russ., abstract in Eng.).
19. Belov, M.A., Grishko, S.I., Lishilin, M.V., Osipov, P.A., Cheremisina, E.N. (2021). Strategy for Training IT Professionals Using the Innovative Training Data Center "Virtual Computer Lab" to Effectively Solve the Problems of Digital Transformation and Acceleration of the Digital Economy. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie = Modern Information Technologies and IT-Education*. Vol. 17, no. 1, pp. 134-144, doi: 10.25559/SITITO.17.202101.703 (In Russ., abstract in Eng.).
20. Naumov, V.N. (2020). Analysis of the Results of the Kaggle Survey on the State of Data Science and Machine Learning in Russia. *Nauchnye trudy Severo-Zapadnogo instituta upravleniya RANKbiGS [Proceedings of the North-Western Institute of Management RANEPa]*. Vol. 11, no. 4 (46), pp. 172-184. Available at: https://spb.ranepa.ru/wp-content/uploads/2021/03/nt_11_4.pdf (accessed 26.03.2022). In Russ., abstract in Eng.).
21. Anuva, S. (2018). Analytical Study of SWAYAM. *IJRAR – International Journal of Research and Analytical Reviews*. Vol. 5, issue 3. Available at: http://ijrar.com/upload_issue/ijrar_issue_1471.pdf (accessed 26.03.2022).

*The paper was submitted 27.02.22
Accepted for publication 26.03.22*