

Образование в информационном обществе

ОБУЧЕНИЕ ДАТА-ГРАМОТНОСТИ: СОДЕРЖАНИЕ И КОНКУРЕНТНЫЕ ПРЕИМУЩЕСТВА ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Статья рекомендована к публикации членом редакционного совета А.М.Елизаровым 15.12.2020.

Дерябин Андрей Александрович

MSc Social Psychology

Научный сотрудник Научно-исследовательского центра социализации и персонализации образования детей Федерального института развития образования Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации

Москва, Россия

deryabin-aa@ranepa.ru

Попов Александр Анатольевич

Доктор философских наук, доцент

Заведующий научно-исследовательским сектором «Открытое образование» Научно-исследовательского центра социализации и персонализации образования детей Федерального института развития образования Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации; заведующий лабораторией компетентных практик образования Института системных проектов Московского городского педагогического университета, профессор кафедры социологии и массовых коммуникаций гуманитарного факультета Новосибирского государственного технического университета

Москва, Россия

popov-aa@ranepa.ru

Аннотация

Представлен анализ зарубежного опыта и основных дискуссионных тем в разработке образовательных модулей науки о данных и машинного обучения для подросткового и юношеского возраста. Особое внимание уделяется таким практическим аспектам разработки содержания обучения как использование аутентичных данных в процессе обучения и обеспечение релевантности данных интересам учащихся. Приводятся характеристики учебных данных с точки зрения их сложности и эффекта на дата-грамотность, критическое мышление и предметные знания учеников. Рассматриваются условия становления субъектности учащихся через культивирование активной позиции «дата-продюсера» в противоположность традиционным дисциплинарным образовательным подходам. Материалом для анализа послужили преимущественно публикации 2014-2020 годов в научных изданиях, индексируемых в базе данных Scopus. Статья подготовлена в рамках выполнения научно-исследовательской работы государственного задания РАНХиГС.

Ключевые слова

дата-грамотность; наука о данных; машинное обучение; образовательные программы; дополнительное образование; анализ данных; data science; data literacy; искусственный интеллект; профориентация; информатика; обществознание

Введение

В контексте фундаментальных изменений в экономике и на рынке труда важное значение приобретает внедрение образовательных программ в области анализа данных и машинного обучения. Поддержание конкурентоспособности экономики требует широкого внедрения технологий искусственного интеллекта и обеспечение этой сферы соответствующими кадрами. Сформировать ответ на вызов сохранения конкурентоспособности призван ряд правительственных инициатив в научно-технологической и образовательной политике, в частности, «Национальная стратегия развития искусственного интеллекта на период до 2030 года». Основными

направлениями повышения уровня обеспечения российского рынка технологий искусственного интеллекта квалифицированными кадрами в ней называются: (а) разработка и внедрение образовательных модулей в рамках образовательных программ всех уровней образования для получения гражданами знаний, приобретения ими компетенций и навыков в области математики, программирования, анализа данных, машинного обучения, способствующих развитию искусственного интеллекта. При этом в целях развития перспективных методов искусственного интеллекта приоритетное значение приобретает конвергентное знание, обеспечиваемое в том числе за счет интеграции математического, естественно-научного и социально-гуманитарного образования.¹ В соответствии с этой повесткой и с учетом динамично развивающейся индустрии данных российские образовательные программы требуют содержательного обновления.

В российском ландшафте школьных дисциплин дата-грамотность и наука о данных как практика аналитического и количественного мышления не имеют пока собственного места. Математика в школе, сосредоточившись на базовой теории, сопротивляется освоению практических методов ее применения в реальной жизни. В школьной физике и биологии данным отводится иллюстративная роль, несмотря на то, что это науки, в которых важную роль играет эксперимент. Общественные дисциплины, всё более дата-ориентированные на практике, едва ли сегодня готовы к тому, чтобы передать школе осмысленное применение данных на уроках обществознания. Пока образовательное сообщество решает задачу, как нужно изменить образовательные системы, чтобы из школы вышли ученики с концептуальным пониманием и навыками анализа данных, наиболее эффективной формой для этого остается дополнительное образование.

Сфера дополнительного образования, в свою очередь, является конкурентной средой, в которой важным критерием востребованности образовательных программ является их содержание и качество. Понятие «дата-грамотность», спектр компетенций в области науки о данных и примеры существующих образовательных программ для учащихся подросткового и юношеского возраста обсуждались нами ранее.² Настоящая статья посвящена тому, как важнейшие аспекты содержания образовательной программы по дата-грамотности, а именно, характер учебных данных и тематика учебных проектов делают обучение более эффективным, придавая образовательному продукту конкурентное преимущество на рынке образовательных услуг.

1 Какие данные использовать для обучения?

С точки зрения ряда специалистов, наибольший учебный эффект на дата-грамотность, критическое мышление и предметные знания старшеклассников имеет погружение в исследование аутентичных научных данных [1]. Авторы широко цитируемого подхода «дата-шагов» Erickson и др. считают, что вводный курс науки о данных должен давать учащимся задачи с разнообразными по содержанию, необработанными, «непослушными» данными [2]. Wolff и др. [49], работая с детьми 10-14 лет, также полагают, что следует предъявлять учащимся не упрощенные, но комплексные данные. Однако «комплексные» данные не обязательно означает «большие», а «реальные» не обязательно означает «подлинные». Чтобы избежать путаницы в терминах, приводим ниже адресованный преподавателям разбор характеристик аутентичных учебных данных, обобщенный в Таблице 1.

Под аутентичными, или подлинными, данными понимаются «достоверные количественные и качественные данные, извлеченные из феноменов реальной жизни». Наоборот, неаутентичные данные могут быть искусственно сгенерированы в демонстрационных целях или могут быть результатом манипулирования данными для достижения задуманного учебного результата или интерпретации. Аутентичные данные могут собраны с использованием различных измерительных методов и инструментов (например, данные сенсоров), сгенерированы компьютерными моделями и симуляциями или получены из онлайн-репозитория и научных публикаций [4, 5, 6]. Kjelvik и Schultheis [1] имеют в виду в первую очередь естественнонаучные данные, однако с некоторыми допущениями и в зависимости от содержания учебной программы преподаватель вполне может отнести к ним неадаптированные экономические, социологические и проч. данные. Сбор, анализ, интерпретация аутентичных датасетов потенциально помогают учащимся улучшить навыки

¹ Указ Президента Российской Федерации от 10.10.2019 г. № 490 О развитии искусственного интеллекта в Российской Федерации [Электронный ресурс] / Сайт Президента РФ. URL: <http://www.kremlin.ru/acts/bank/44731> (дата обращения 10.04.2020).

² Дерябин А.А., Попов А.А. Дата-грамотность как новая цифровая компетенция // Информационное общество, 2020, №5. С. 39-47.

критического мышления, углубить предметные знания и связать свои количественные аналитические навыки со сферой науки о данных [7, 8].³ Ниже представлены характеристики сложности учебных данных, которые могут варьироваться в зависимости от задач учебной программы и подготовленности аудитории.

Охват (scope) данных определяется числом переменных и числом значений, которые принимает каждая переменная. Простые датасеты состоят из небольшого числа переменных и содержат только ту информацию, которая нужна для ответа на вопросы исследования. Сложные, с широким охватом датасеты могут содержать много переменных, как нужных, так и не нужных для решения поставленной исследовательской задачи. «Узкие» датасеты хороши для наглядной демонстрации зависимых и независимых переменных и отношений между ними; работа с «широкими» датасетами может представлять собой дополнительный вызов для учащихся, которые должны определить, какие переменные им нужны, а какие нет, и может стимулировать их исследовательскую активность, которая может вести к творческой переформулировке поставленных задач и выявлению непредвиденных связей между переменными.

Селективность (selection) касается принятия решений о том, какие переменные нужно определить, какие данные, откуда и при помощи каких инструментов необходимо собрать для решения поставленной исследовательской задачи. В простейшем случае переменные задаются преподавателем. В случае, если учащимся предлагается сделать это самостоятельно, им, тем самым, предоставляется возможность самим определить свою образовательную траекторию [9]. Последнее, разумеется, обеспечивает более богатый образовательный и исследовательский опыт [10, 8, 11].

Курирование (curation) данных – это организация и интеграция данных, собранных из разных источников, подготовка их к анализу и обеспечение их целостности в этом процессе; курирование может включать в себя обнаружение и сбор данных, аннотирование, подготовку и очистку, представление, организацию доступа к данным и их хранение. Преподаватель зачастую обеспечивает учащихся хорошо организованными, готовыми к использованию датасетами, чтобы облегчить их задачу. Однако, когда курирование делегируется студентам, у них появляется возможность освоить новые навыки работы с данными [10]: синтез нескольких датасетов в один предполагает приведение их к единому формату и размерности, кодирования категориальных переменных в численные и проч., что требует умений работы с электронными таблицами и соответствующими программными методами.

Беспорядок (messiness) – наличие в датасете пропусков и выбросов, случайных и систематических ошибок, влияющих на изменчивость переменных. Подготовка сырых данных – обычно едва ли не самый длительный этап в цикле промышленного дата-проекта, и навыки чистки «грязных» данных имеют большую ценность как для лучшего понимания учащимися источников изменчивости признаков в датасете, так и для приобретения практических навыков манипулирования данными.

Размер (size) данных определяется числом переменных и числом значений переменной (например, в случае табличных данных, это число столбцов и строк соответственно). В учебном процессе размер учебного датасета может варьироваться от маленького на начальных этапах обучения, до большого в ходе выполнения студенческого проекта.

Обеспечение учащихся опытом работы с разными типами данных вместе с обсуждением контекста использования этих данных и того, как именно эти данные помогают ответить на вопросы их исследования, может быть важным для развития дата-грамотности [12]. Ряд педагогических инициатив позволяет предположить, что взаимодействие с аутентичными данными положительно влияет на дата-грамотность учащихся [9, 13]. Вместе с тем, недостаток исследований методик, использующих аутентичные данные, не позволяет сделать вывод об однозначной эффективности этого подхода [14]. Исследований, сравнивающих эффекты аутентичных и неаутентичных данных на результаты обучения в настоящее время нами не обнаружено. Чтобы продемонстрировать эту зависимость, требуются исследования, которые бы изолировали воздействие аутентичных данных на дата-грамотность, интересы учащихся и их вовлеченность в изучение данных через манипулирование использованием аутентичных и неаутентичных данных в процессе обучения.

³ Цит. по [1].

Таблица 1. Характеристики сложности учебных данных по Kjelvik и Schultheis [1].

Характеристика	← Простота		Сложность →	
Охват (scope)	Узкий: есть только нужные данные		Широкий: есть нужные и ненужные данные	
Селективность (selection)	Переменные заданы преподавателем	Студенты выбирают переменные из предложенного датасета	Студенты самостоятельно собирают датасет с нужными им переменными	
Курирование (curation)	Полное: датасет дается студентам в готовом к анализу виде	Частичное: сырые данные, готовые к анализу, но не аннотированные	Отсутствует: студенты должны аннотировать сырые данные и подготовить их к анализу, трансформировать	Синтез: студенты должны объединить несколько датасетов и подготовить их к анализу
Размер (size)	Маленький: несколько переменных и их значений		Большой: много переменных и их значений	
Беспорядок (messiness)	Чистые: нет пропусков данных и выбросов, низкая изменчивость		Грязные: есть пустые значения и выбросы, высокая изменчивость	

2 Релевантность данных интересам учащихся как условие эффективной реализации образовательной задачи

Для эффективного обучения данные и результаты работы с ними должны быть «присвоены» студентами, а для этого они должны быть не только аутентичными, но и релевантными интересам учащихся. Когда учащиеся не представляют, как данные, с которыми им предложили работать, связаны с реальной жизнью, учебные задачи и действия могут представляться им затруднительными или бессмысленными [1]. Некоторые могут испытывать трудности с интерпретацией неаутентичных данных [15, 16] или с применением новых умений в других контекстах [17].

Напротив, аутентичные данные, связанные с предметным содержанием, с понятным контекстом их создания и в ясной связи с проблемами реального мира потенциально интересней и привлекательней для учащихся [12, 18], и от них поступает обратная связь, которая говорит об эмоциональной вовлеченности в работу с такими данными [8, 19].

Как обеспечить это присвоение данных в образовательном процессе? Через релевантную интересам учеников тематику учебных задач, сопутствующих им данных и взаимодействие с ними. Образовательная программа должна предусматривать решение таких образовательных задач, которые, с одной стороны, имеют статус «настоящих» (то есть актуальных и практически значимых) для всех ее участников, а с другой стороны, задают некоторую неопределенность как в ее интерпретации, так и в сценариях их решения.⁴ Это может быть привнесено в образовательный процесс как за счет самостоятельного определения учащимися целей и задач их дата-проекта, так и за счет использования необработанных аутентичных данных. Такая специфика программы позволяет обеспечить продуктивную деятельность учащихся и удерживать их высокую мотивацию на использование данных и техник работы с данными как с образовательным ресурсом.

Wilkerson и Polman [20] отмечают, что изучение науки о данных – это не просто утилитарный процесс предметного обучения, но выстраивание многослойных отношений учащегося по отношению к данным как социокультурному тексту: каждый учащийся может обнаружить данные, которые рассказывают о нем что-то важное, будь то данные от сенсоров носимых им «умных часов», геометки в истории его навигатора или обобщенная социально-экономическая статистика о регионе его проживания.

Lee и Dubovi [21] показывают, как возникает глубокая и сложная связь с данными у семей с диабетом 1 типа в процессе использования ими устройств мониторинга и постоянных измерений. Kahn [22] описывает конструирование «семейных геобиографий» на основе данных и

⁴ Глухов П.П. Компетентностные испытания как современная форма оценки образовательных достижений // Философия образования, 2016, № 4 (67). С. 99-110.

демонстрирует, как их интерпретации в сочетании с семейными историями миграции формируют вокруг них многослойный социоисторический контекст, влияющий на молодое поколение.

Кроме того, для создания условий эффективной профессиональной пробы важно дать учащимся попробовать себя в разных позициях в ходе решения задач по отношению к данным – например, в разных ролях в соответствии с разделением труда внутри дата-проекта (программист, дата-инженер, менеджер проекта и предметный эксперт). Так, Wise [23], описывая свой проект, в котором студенты собирают данные о собственном обучении, выделяет три взаимодополняющих позиции по отношению к данным, между которыми переключаются ее студенты в ходе учебного проекта: «внутренняя» (респондент, источник данных), «внешняя» (реципиенты, которым адресован аналитический отчет о проделанной работе, например «заказчик») и «аналитик данных» (тот, кто работает с данными).

Hardy [24] подчеркивает важность становления субъектности учащихся в генерации знания через культивирование активной позиции «дата-продюсера» в противоположность традиционным, дисциплинарным образовательным подходам, нацеленным на воспроизводство заранее известного и дидактического «правильного» результата и «правильных» данных, которые словно бы ждут, чтобы их собрали.

3 Метапредметные компетенции: «думать данными»

Говоря о профориентационной компоненте преподавания науки о данных в подростковом и юношеском возрасте, следует учитывать актуальные требования, предъявляемые к современному специалисту: не только соотносить конкретную задачу с перечнем испытанных вариантов решения и выбирать из этого перечня оптимальный вариант, но и конструировать новые способы решения, оптимальные именно для данного конкретного случая; оформлять проблемную ситуацию и ставить задачи, исходя из неё; намечать необходимый результат для своих действий в рамках конкретного проекта или технологического процесса, и выстраивать программу действий по достижению этого результата. Для этого у специалиста должен быть сформирован тип мышления, соответствующий выбранной им сфере профессиональной деятельности, позволяющий понимать присущие этой сфере технологии, приёмы работы и тип результата в их взаимосвязи, как динамическую систему, способную меняться и развиваться как вследствие обновления технологий, так и вследствие изменения требований к конечному продукту.⁵

Так, Суосо [25] пишет о важности формирования «ментальной привычки» как способа мышления, вопрошания, и решения задач. Finzer [26] также приверженец идеи, что в обучении в школе нужно прививать специфическое мышление, взгляд на мир сквозь призму данных. Это видение прослеживается и у авторов, пишущих о разработке образовательных программ по Data Science и статистике на уровне высшей школы [27, 28].

Заключение

В современных формах инженерного образования принципиальным становится требование к студенту: для того, чтобы профессиональное знание было эффективно освоено и присвоено, уже на старте получения профессионального образования студент должен иметь образ будущей профессии, понимать, каким образом осваиваемые знания и способы усиливают его как будущего специалиста. Следовательно, работа по формированию у будущих дата-аналитиков специфически профессионального мышления и основ профессиональной культуры должна происходить ещё в период их обучения в школе, в режиме пропедевтики будущей деятельности.

Формирование образа будущей профессии дата-аналитика и связанного с ней типа мышления может обеспечиваться практико-ориентированным образованием, где в качестве единицы при построении учебного процесса используется реально существующая проблема, которая переживается учеником как значимая.⁶ Эта проблема может быть описана посредством аутентичных количественных данных и оформлена в виде образовательной задачи, организующей одновременно пробное практическое действие ученика и его рефлексии в отношении используемых им инструментов дата-аналитики.

⁵ Попов А.А. Образовательные программы и элективные курсы компетентностного подхода / Изд. 6-е. – М.: Ленанд, 2019. – С.170.

⁶ Глухов П.П. Практическое мышление как компетентностный образовательный результат / Социальные коммуникации и эволюция обществ. Материалы V Международной научно-практической конференции. Под редакцией И.А. Вальдмана; Новосибирск: 2016. С. 447-453.

Литература

1. Kjelvik M. K., Schultheis, E. H. Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. // CBE Life Sciences Education, 2019, 18(2). С. 1–8.
2. Erickson T. и др. Data Moves: one key to data science at school level / Proceedings of the International Conference on Teaching Statistics (ICOTS-10), 2018. [Электронный ресурс]. URL https://iase-web.org/Conference_Proceedings.php?p=ICOTS_10_2018 (дата обращения 10.04.2020).
3. Wolff A., Wermelinger M., Petre M. Exploring design principles for data literacy activities to support children's inquiries from complex data // International Journal of Human Computer Studies, 2019, 129(March). С. 41–54. doi:10.1016/j.ijhcs.2019.03.006
4. Magnusson S. J. и др. How should learning be structured in inquiry-based science instruction? Investigating the interplay of 1st- and 2nd-hand investigations / In Proceedings of the 6th international conference on learning sciences. Alpharetta, GA: International Society of the Learning Sciences, 2004. С. 318–325.
5. Hug B., McNeill K. L. Use of first-hand and second-hand data in science: Does data type influence classroom conversations? // International Journal of Science Education, 2008, 30. С. 1725–1751.
6. Kerlin, S. C., McDonald, S. P., & Kelly, G. J. Complexity of secondary scientific data sources and students' argumentative discourse // International Journal of Science Education, 2010, 32(9). С. 1207–1225.
7. Mourad T., Grant B. W., Gram W. K. Engaging undergraduate students in ecological investigations using large, public datasets // Teaching Issues and Experiments in Ecology, 2012, 8.
8. Langen T. A. и др. Using large public datasets in the undergraduate ecology classroom // Frontiers in Ecology and the Environment, 2014, 12. С. 362–363.
9. Gould R., Sunbury S., Dussault M. (2014). In praise of messy data // Science Teacher, 2014, 81(8). С. 31.
10. Carlson J. R. и др. Determining data information literacy needs: A study of students and research faculty. [Электронный ресурс]. // Libraries Faculty and Staff Scholarship and Research, Paper 23. URL https://docs.lib.purdue.edu/lib_fsdocs/23/ (дата обращения 10.04.2020)
11. Calzada Prado, J., Marzal, M. Incorporating data literacy into information literacy programs: Core competencies and contents // Libri, 2013, 63(2). С. 123–134.
12. Schultheis E. H., Kjelvik, M. K. Data Nuggets: Bringing real data into the classroom to unearth students' quantitative and inquiry skills // American Biology Teacher, 2015, 77(1). С. 19–29.
13. Kastens K. A., Krumhansel R., Baker I. Thinking big – Transitioning your students from working with small, student-collected data sets towards “big data.” // Science Teacher, 2015, 82(5). С. 25–31.
14. Aikens M. L., Dolan E. L. Teaching quantitative biology: Goals, assessments, and resources // Molecular Biology of the Cell, 2014, 25. С. 3478–3481.
15. Piatek-Jimenez K. и др. Helping students become quantitatively literate. / Mathematics Teacher, 105(9), 2012. С. 692–696.
16. Common Core State Standards for Mathematics [Электронный ресурс]. // Common Core State Standards Initiative (CCSSI), 2014. URL <http://www.corestandards.org/Math> (дата обращения 10.04.2020)
17. Borges-Rey E. L. Data literacy and citizenship: Understanding “big data” to boost teaching and learning in science and mathematics / Ramírez-Montoya M.-S. (Ed.), Handbook of research on driving STEM learning with educational technologies. Hershey, PA: IGI Global, 2017. - С. 65–79.
18. Hulleman C. S., Harackiewicz J. M. Promoting interest and performance in high school science classes // Science, 2009, 326(5958). С. 1410–1412.
19. Wolff и др. Creating an Understanding of Data Literacy for a Data-driven Society [Электронный ресурс]. // The Journal of Community Informatics, 2016, 12(3). URL <http://ci-journal.net/index.php/ciej/article/view/1286> (дата обращения 10.04.2020).
20. Wilkerson M. H., Polman J. L. Situating Data Science: Exploring How Relationships to Data Shape Learning. Journal of the Learning Sciences, 2020, 29(1). С. 1–10. doi:10.1080/10508406.2019.1705664
21. Lee V. R., Dubovi I. At home with data: Family engagements with data involved in type 1 diabetes management // Journal of the Learning Sciences, 2020, 29(1). С. 11–31. doi:10.1080/10508406.2019.1666011

22. Kahn J. Learning at the Intersection of Self and Society: The Family Geobiography as a Context for Data Science Education // *Journal of the Learning Sciences*, 2020, 29 (1), С. 57-80.
23. Wise A. F. Educating Data Scientists and Data Literate Citizens for a New Generation of Data // *Journal of the Learning Sciences*, 2020, 29(1). С. 165–181.
24. Hardy L., Dixon C., Hsi S. From Data Collectors to Data Producers: Shifting Students' Relationship to Data // *Journal of the Learning Sciences*, 2020, 29(1). С. 104–126.
25. Cuoco A., Goldenberg E. P., Mark J. Habits of Mind: an organizing principle for mathematics curriculum // *Journal of Mathematical Behavior*, 1997, 15(4). С. 375-402.
26. Finzer, W. The data science education dilemma [Электронный ресурс]. // *Technology Innovations in Statistics Education*, 2013, 7(2). URL <https://escholarship.org/uc/item/7gv0q9dc> (дата обращения 10.04.2020).
27. De Veaux, R. и др. Curriculum guidelines for undergraduate programs in data science // *Annual Review of Statistics and Its Application*, 2017, 4(1). С. 15–30.
28. Hardin J. и др. Data science in statistics curricula: Preparing students to “think with data” // *The American Statistician*, 2015, vol. 69(4). С. 343-353.

TEACHING DATA LITERACY: CURRICULUM CONTENT AND COMPETITIVE ADVANTAGES

Deryabin, Andrey Aleksandrovich

MSc Social Psychology

Researcher at Federal Education Development Institute, Russian Presidential Academy of National Economy and Public Administration

Moscow, Russia

deryabin-aa@ranepa.ru

Popov, Aleksandr Anatolyevich

Dr. Sci. (Philos.), Associate Professor

Head of Open Education section at Federal Education Development Institute, Russian Presidential Academy of National Economy and Public Administration; Head of Competency Practices Lab, Institute of System Projects, Moscow City University; Professor of Sociology & Mass Communications Dept., Novosibirsk State Technical University

Moscow, Russia

popov-aa@ranepa.ru

Abstract

A review of publications and an analysis of major discussions about the Data Science curriculum development for middle and secondary school is presented with special emphasis on use of authentic data and data relevancy to students' context.

Keywords

big data; data literacy; data science education; machine learning; artificial intelligence; school education; computer science education; curriculum development

References

1. Kjeldvik M. K., Schultheis, E. H. Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. // CBE Life Sciences Education, 2019, 18(2). С. 1-8.
2. Erickson T. и др. Data Moves: one key to data science at school level / Proceedings of the International Conference on Teaching Statistics (ICOTS-10), 2018. [Электронный ресурс]. URL https://iase-web.org/Conference_Proceedings.php?p=ICOTS_10_2018 (дата обращения 10.04.2020).
3. Wolff A., Wermelinger M., Petre M. Exploring design principles for data literacy activities to support children's inquiries from complex data // International Journal of Human Computer Studies, 2019, 129(March). С. 41-54. doi:10.1016/j.ijhcs.2019.03.006
4. Magnusson S. J. и др. How should learning be structured in inquiry-based science instruction? Investigating the interplay of 1st- and 2nd-hand investigations / In Proceedings of the 6th international conference on learning sciences. Alpharetta, GA: International Society of the Learning Sciences, 2004. С. 318-325.
5. Hug B., McNeill K. L. Use of first-hand and second-hand data in science: Does data type influence classroom conversations? // International Journal of Science Education, 2008, 30. С. 1725-1751.
6. Kerlin, S. C., McDonald, S. P., & Kelly, G. J. Complexity of secondary scientific data sources and students' argumentative discourse // International Journal of Science Education, 2010, 32(9). С. 1207-1225.
7. Mourad T., Grant B. W., Gram W. K. Engaging undergraduate students in ecological investigations using large, public datasets // Teaching Issues and Experiments in Ecology, 2012, 8.
8. Langen T. A. и др. Using large public datasets in the undergraduate ecology classroom // Frontiers in Ecology and the Environment, 2014, 12. С. 362-363.
9. Gould R., Sunbury S., Dussault M. (2014). In praise of messy data // Science Teacher, 2014, 81(8). С. 31.

10. Carlson J. R. и др. Determining data information literacy needs: A study of students and research faculty. [Электронный ресурс]. // Libraries Faculty and Staff Scholarship and Research, Paper 23. URL https://docs.lib.purdue.edu/lib_fsdocs/23/ (дата обращения 10.04.2020)
11. Calzada Prado, J., Marzal, M. Incorporating data literacy into information literacy programs: Core competencies and contents // Libri, 2013, 63(2). С. 123–134.
12. Schultheis E. H., Kjellvik, M. K. Data Nuggets: Bringing real data into the classroom to unearth students' quantitative and inquiry skills // American Biology Teacher, 2015, 77(1). С. 19–29.
13. Kastens K. A., Krumhansel R., Baker I. Thinking big – Transitioning your students from working with small, student-collected data sets towards “big data.” // Science Teacher, 2015, 82(5). С. 25–31.
14. Aikens M. L., Dolan E. L. Teaching quantitative biology: Goals, assessments, and resources // Molecular Biology of the Cell, 2014, 25. С. 3478–3481.
15. Piatek-Jimenez K. и др. Helping students become quantitatively literate. / Mathematics Teacher, 105(9), 2012. С. 692–696.
16. Common Core State Standards for Mathematics [Электронный ресурс]. // Common Core State Standards Initiative (CCSSI), 2014. URL <http://www.corestandards.org/Math> (дата обращения 10.04.2020)
17. Borges-Rey E. L. Data literacy and citizenship: Understanding “big data” to boost teaching and learning in science and mathematics / Ramirez-Montoya M.-S. (Ed.), Handbook of research on driving STEM learning with educational technologies. Hershey, PA: IGI Global, 2017. - С. 65–79.
18. Hulleman C. S., Harackiewicz J. M. Promoting interest and performance in high school science classes // Science, 2009, 326(5958). С. 1410–1412.
19. Wolff и др. Creating an Understanding of Data Literacy for a Data-driven Society [Электронный ресурс]. // The Journal of Community Informatics, 2016, 12(3). URL <http://ci-journal.net/index.php/ciej/article/view/1286> (дата обращения 10.04.2020).
20. Wilkerson M. H., Polman J. L. Situating Data Science: Exploring How Relationships to Data Shape Learning. Journal of the Learning Sciences, 2020, 29(1). С. 1–10.
doi:10.1080/10508406.2019.1705664
21. Lee V. R., Dubovi I. At home with data: Family engagements with data involved in type 1 diabetes management // Journal of the Learning Sciences, 2020, 29(1). С. 11–31.
doi:10.1080/10508406.2019.1666011
22. Kahn J. Learning at the Intersection of Self and Society: The Family Geobiography as a Context for Data Science Education // Journal of the Learning Sciences, 2020, 29 (1), С. 57–80.
23. Wise A. F. Educating Data Scientists and Data Literate Citizens for a New Generation of Data // Journal of the Learning Sciences, 2020, 29(1). С. 165–181.
24. Hardy L., Dixon C., Hsi S. From Data Collectors to Data Producers: Shifting Students' Relationship to Data // Journal of the Learning Sciences, 2020, 29(1). С. 104–126.
25. Cuoco A., Goldenberg E. P., Mark J. Habits of Mind: an organizing principle for mathematics curriculum // Journal of Mathematical Behavior, 1997, 15(4). С. 375–402.
26. Finzer, W. The data science education dilemma [Электронный ресурс]. // Technology Innovations in Statistics Education, 2013, 7(2). URL <https://escholarship.org/uc/item/7gv0q9dc> (дата обращения 10.04.2020).
27. De Veaux, R. и др. Curriculum guidelines for undergraduate programs in data science // Annual Review of Statistics and Its Application, 2017, 4(1). С. 15–30.
28. Hardin J. и др. Data science in statistics curricula: Preparing students to “think with data” // The American Statistician, 2015, vol. 69(4). С. 343–353.