# Дата-грамотность как актуальный образовательный результат в условиях цифровой трансформации общества

Павел Павлович Глухов [1], Андрей Александрович Дерябин [2], and Александр Анатольевич Попов[1]
[1]ГАОУ ВО МГПУ, Москва, Россия; РАНХиГС, Москва, Россия
2 РАНХиГС, Москва, Россия

**Аннотация.** *Наука о данных (Data Science) затрагивает все больший круг сфер обыденной жизни, но общее образование в России пока еще не отреагировало на новые вызовы, связанные с этим аспектом цифровизации. Изменения в технологиях, экономике и обществе за последние два десятилетия сформировали новую повестку как для преподавания математики и информационных технологий, так и для общественных наук и медиа-образования. Образование во всех этих областях требует переосмысления содержания и методов преподавания в связи с растущим значением науки о данных и искусственного интеллекта в контексте фундаментальных изменений в экономике и на рынке труда. Так как многие сферы человеческой жизни сегодня претерпевают изменения, то возникает необходимость оформления новых типов и видов образовательных результатов, на которые должна быть направлена современная педагогика. Один из современных способов ответить на подобные вызовы – это выделение новых грамотностей (медиа-грамотность, экологическая грамотность, функциональная грамотность и т.д.). В статье рассматривается такое понятие как дата-грамотность, исследуется его содержание и состав, обосновывается её актуальность как образовательного результата, сообразного трендам цифровизации, которые мы можем фиксировать в современном обществе. Осуществляется различение подходов к углубленному и общему изучению науки о данных. Даётся описание различных типов задач, направленных на развитие дата-грамотности у учащихся в условиях их постановки на разном учебном материале. Рассматриваются возможные способы разворачивания программ, направленных на освоение учащимися науки о данных без необходимости её оформления в отдельную дисциплину или школьный предмет.*

**Ключевые слова:** *наука о данных, новые грамотности, дата-грамотность, образование.*

## Введение

Сегодня мы живем в цифровом обществе, одной из важнейших характеристик которого является «датафицированность» - когда многие процессы экономического и социального развития основываются на сборе и обработке больших массивов данных (Big-Data[2]), в том числе персональных. В определённом смысле современный человек становится источником данных, которые выражаются в том, что он покупает, по каким ссылкам переходит в сети Интернет, сколько времени затрачивает на просмотр той или иной информации, в каких геолокациях бывает и т.д. Становясь обладателями того или иного информационного продукта (смартфона, приложения или иного программного обеспечения) мы вступаем в отношение с разными алгоритмами, собирающими информацию о наших действиях и интересах. Встаёт вопрос об информационной безопасности и степенью её осознанности конкретными людьми. Когда мы обсуждаем молодых людей, берущих в руки смартфон в достаточно раннем возрасте, данный вопрос встаёт ещё более актуальным. Однако мы не должны говорить об изоляции молодых людей от современных достижений в области информационных технологий. Формирование специфических типов грамотностей видится как более рациональная задача, которая как минимум позволяет повысить уровень

[*] Corresponding author: sjelene@yandex.ru

[2] Big-Data - данные, объем которых, многообразие и скорость прироста превосходят возможности традиционных систем хранения и обработки информации, существовавших на рубеже 2000-х годов

осознанности, критичности и рефлексии по отношению к процессам эксплуатации современных технических средств и информации. Целью исследования является определение содержания специфического типа грамотности как возможного нового образовательного результата, ориентированного на рефлексивное отношение к феномену больших данных в современном обществе.

## Методы

Исследование носит обзорный характер, где коллектив авторов попытался проанализировать и обобщить современные концепты образовательных результатов, связанных с феноменом больших данных. Подбирались те источники, которые позволяют рассматривать грамотность в области Big-Data с позиции общеразвивающих, а не (пред-) профессиональных образовательных результатов.

## Результаты

Несмотря на то, что анализ больших массивов данных чаще относят к узкоспециализированным и профессиональным областям деятельности, инструменты науки о данных (Data Science) представляются более массово доступными, чем может показаться на первый взгляд. Сегодня обретает чёткие границы такое понятие как дата-грамотность (data literacy), которую можно описать как способность общего характера, применимой в рамках общественной и этической проблематики, позволяющей вырабатывать обоснованные рациональные решения и оформлять доказательную позицию в современном цифровом мире (M. Schield, 2018). Pangrazio и Sefton-Green справедливо подчеркивают, что к данным следует относиться критически и с осторожностью, ведь используемые датасеты (систематизированные наборы данных) могут претерпевать слишком много циклов обработки и использоваться разными агентами в разных контекстах, с разными целями и мотивами. Дата-грамотность предполагает владение инструментами количественного и качественного анализа данных, но гораздо более важным умением здесь выступает применение подобных инструментов в определённом контексте (J. P. Gibson, T. Mourad, 2018).

Дата-грамотность представляется как комплекс разных характеристик современного члена информационного общества и может включать в себя конфигурации следующих способностей:

- оценка последствий использования и предоставления данных о себе и своей жизни при работе с современными мобильными устройствами, платёжными системами, социальными сетями, разнообразными информационными сервисами и программным обеспечением и т.д.;
- реконструкция контекста кем-либо систематизированной преподносимой информации до уровня используемых в ней данных и методов их применения;
- поиск, обнаружение, извлечение и обработка данных, необходимых для более рационального принятия жизненных решений, которые могут быть связаны с выбором города проживания, места работы или с прогнозированием финансового положения собственной семьи.

## Обсуждение

Один из ключевых вопросов, связанных с дата-грамотностью – это избыточность и целесообразность использования методов и средств обучению науке о данных. Что мы можем упразднить в классических способах изучения науки о данных для того, чтобы выйти на уровень формирования общеразвивающих образовательных результатов, находящих широкое применение на разных (в т.ч. и ранних) этапах развития?

Чаще всего мы можем встретить широкий спектр предложений именно в области специализированных образовательных программ, посвящённых Data Science. Широко

распространены онлайн-курсы, которые пытаются реагировать на кадровый заказ[3]. Также мы можем увидеть большое количество программ профессиональной подготовки в области Data Science на уровне бакалавриата и магистратуры (R. De Veaux, 2017; B. S. Baumer, 2015; P. Anderson, 2014). Такие программы сильно акцентирую внимание на технической стороне процесса анализа данных, что заставляет смотреть на них как на программы узкоспециализированной подготовки. Такой подход может быть избыточным для решения задач массового формирования интересующего нас типа грамотности, ведь в таком случае мы уделяем меньше внимания общему влиянию данных на жизнь общества, экономики и культуры, что также отмечается экспертами (B. Heinemann, 2018). В свою очередь те примеры, которые затрагивают общий контур обучения науке о данных детей подросткового и юношеского возраста имеют разную продолжительность, а также в них всех присутствует разброс содержания (S. Srikant & V. Aggarwal, 2017; C. Bryant, 2019; A. Dryer, N. Walia, A. Chattopadhyay, 2018; S. Datta, V. Nagabandi, 2018; R. Mariescu-Istodor, I. Jormanainen, 2019; A. Wolff, M. Wermelinger, M. Petre, 2019).

Как утверждает ряд специалистов по обучению науке о данных, для успешного освоения этой области важны не столько сложные познания в области специализированного программного обеспечения, сколько сочетания базисных познаний в этой области со способностью к анализу[4]. Зависимость всё больших сфер нашей жизни от больших массивов данных актуализирует данные характеристики личности в общеразвивающем контексте. В этом моменте мы можем обсуждать не столько возможность формирования навыков работы с большими данными, сколько общих аналитических навыков, позволяющих человеку критически относиться к информации, которая использует большие данные или массивы специализированных данных. Это может быть связано с ростом такого явления как спекуляция на сложности ради сомнительного обоснования каких-либо социальных и экономических явлений за счёт манипуляции большими массивами данных или корпусами статистических данных. Мы увидели яркую демонстрацию подобного явления в ходе распространения противоречивой информации о распространении COVID-19. Люди склонны доверять графикам и статистикам, т.к. они наукообразны и оперируют фактическими цифрами. Но несмотря на наличие цифр, отражающих факты, они могут быть применены некорректно (например, применяемые данные могут быть добыты некорректными методами или могли проходить множество циклов отчистки и т.д.). В этом смысле обучение дата-грамотности – это в том числе процесс формирования компетентной критичности, которая позволяет не только усомниться в достоверности информации, но также проверить достоверность данных, лежащих в основе предоставляемой нам информации.

Cuoco и др. (A. Cuoco, E. P. Goldenberg, J. Mark, 1997) пишут о важности формирования “ментальной привычки” обращения к данным как способа мышления, вопрошания, и решения задач. Finzer (W. Finzer, 2013) также приверженец идеи, что в обучении в школе нужно прививать специфическое мышление, взгляд на мир сквозь призму данных. Он замечает, что «думать данными» - это своего рода мета-навык, и его развитие в школе должно происходить через применение методов науки о данных в рамках существующих предметов, а не через введение специальной дисциплины. Это видение прослеживается и у авторов, пишущих о разработке образовательных программ по Data Science и статистике на уровне высшей школы (R. De Veaux, 2017; J. Hardin, R. Hoerl, 2015). При этом, на сегодняшний день большие данные и наука о данных практически не выделяются из всего массива дистанционных цифровых технологий в образовании. В учебные планы предметов «Информатика» и «Технология» они входят как частные компоненты базового учебного плана, преподаваемые в отстраненном, объективированном режиме. Редко компьютерный анализ данных и машинное обучение позиционируются для учеников в качестве

[3] Примеры курсов на Coursera https://www.coursera.org/browse/data-science и edX https://www.edx.org/course/subject/data-science

[4] Data science — наука о данных. Как стать data scientist с нуля // [Электронный ресурс] URL: https://future2day.ru/data-science/ (дата обращения 25.12.2019)

инструмента для решения учебных задач, в том числе, проектно-творческого характера (Мамедова Г.А., Зейналова Л.А., Меликова Р.Т., 2017).

## Заключение

Как можно увидеть, дата-грамотность связана не столько с Big-Data, сколько с данными вообще, методами их обработки и аналитическими способностями общего спектра, позволяющими реконструировать основу влияющей на нас информации и формировать валидную основу для принятия собственных рациональных решений. Стоит также понимать, что дата-грамотность отличается от цифровой и информационной грамотности, но соотносится с ними, определённым образом их конкретизируя и дополняя. Если цифровая грамотность позволяет человеку эффективно и безопасно применять цифровые технологии и ресурсы, использовать их сообразно контексту поставленных задач, то дата-грамотность позволяет увидеть и понять какие данные формируются и используются в процессе взаимодействия с цифровыми технологиями. В зависимости от нашего поведения при взаимодействии с тем или иным сервисом мы можем получать разный результат, который и основан на тех типах данных, которые собираются при таком взаимодействии. Например, стоимость поездки на такси, которое мы заказываем при помощи приложений, зависит от разных переменных, в том числе от частоты использования приложения, от истории наших заказов, от истории наших отказов от поездок, от уже проявленной готовности платить больше или меньше в ситуациях пикового спроса на такси и т.д. В зависимости от персональных наборов подобных данных, которые собираются на протяжении всего процесса взаимодействия пользователя с программой, стоимость идентичных поездок может существенно различаться для двух разных пользователей. Важно понимать методы и логику обработки данных учитывая, что они меняются в зависимости от того, сервисы какой компании их используют и для каких задач.

Информационная грамотность формируется вокруг способностей критически относиться к окружающей нас информации, рефлексивно искать и применять информацию для решения тех или иных задач. С одной стороны, дата-грамотность позволяет установить, что лежит в основе той или иной сложноорганизованной информации, например, использующей статистику. В таких случаях дата-грамотность позволяет оценить корректность и избыточность используемой статистики. С другой стороны, дата-грамотность позволяет реконструировать наборы сведений, лежащих в основе алгоритмов, которые рекомендуют нам ту или иную информацию или рекламу. Демонстративной здесь выступает прецедент, связанный с компанией «Cambridge Analytica», которая собрала персональные данные миллионов пользователей социальной сети «Facebook» для моделирования электорального поведения с целью дальнейшего подбора релевантной политической рекламы, влияющей на конечные политические предпочтения категории «неопределившихся» избирателей. Можно предположить, что у человека, который разбирается в методах сбора и обработки данных возникнут подозрения в случае наблюдаемого роста рекомендуемого контента, связанного с конкретной политической партией.

Применение науки о данных в области образования уже позволяет на настоящий момент оформить два разных образовательных результата, на которые она может работать: общеразвивающий и (пред-) профессиональный. Дата-грамотность стоит рассматривать как мета-навык, формируемый соответствующими подходами, не выделяя его в качестве отдельной дисциплины или предмета, если только мы не ставим перед собой целью осуществлять профессиональную или узкоспециализированную подготовку. Дата-грамотность может рассматриваться как самодостаточный образовательный результат, но интегрированный в иные виды и типы учебной деятельности.

## Литература

1. *Мамедова Г.А., Зейналова Л.А., Меликова Р.Т.* Технологии больших данных в электронном образовании // Открытое образование. 2017. – С. 41-48.

2. *Anderson P. и др.* An undergraduate degree in data science: Curriculum and a decade of implementation experience / SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education, 2014. 145–150.

3. *Baumer B. S.* A data science course for undergraduates: Thinking with data // The American Statistician, 2015, 69. C. 334–342.

4. *Bryant C. и др.* A middle-school camp emphasizing data science and computing for social good // SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education, 2019. C. 358–364. doi:10.1145/3287324.3287510

5. *Cuoco, A., Goldenberg, E. P., Mark, J.* Habits of Mind: an organizing principle for mathematics curriculum. // Journal of Mathematical Behavior, vol. 15, no. 4, pp. 375-402, 1997.

6. *Datta S., Nagabandi V.* (2018). Integrating data science and R programming at an early stage / IEEE 4th International Conference on Soft Computing and Machine Intelligence, ISCMI 2017, 2018. C. 1–5. doi:10.1109/ISCMI.2017.8279587

7. *De Veaux, R. и др.* Curriculum guidelines for undergraduate programs in data science // Annual Review of Statistics and Its Application, 2017, 4(1). C. 15–30.

8. *Dryer A., Walia N., Chattopadhyay A.* A middle-school module for introducing data-mining, big-data, ethics and privacy using rapidminer and a Hollywood theme / SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education, 2018. C. 753–758. doi:10.1145/3159450.3159553

9. *Finzer, W.* The data science education dilemma [Электронный ресурс] // Technology Innovations in Statistics Education, vol. 7, no. 2, p. URL https://escholarship.org/uc/item/7gv0q9dc (дата обращения 10.04.2020), 2013.

10. *Gibson J. P., Mourad T.* The growing importance of data literacy in life science education. American Journal of Botany, 105(12), 1–4, 2018.

11. *Hardin, J., Hoerl, R., и др.* Data Science in Statistics Curricula: Preparing Students to “Think with Data” // The American Statistician, p. doi: 10.1080/00031305.2015.1077729, 2015.

12. *Heinemann B. и др.* Drafting a data science curriculum for secondary schools / ACM International Conference Proceeding Series, 2018. C. 1–5. doi:10.1145/3279720.3279737

13. *Mariescu-Istodor R., Jormanainen I.* Machine Learning Exercises for High School Students [Электронный ресурс]. / Proceedings of the 19th Koli Calling International Conference on Computing Education Research, 2019. URL http://www.cs.columbia.edu/~CS4HS/talks/ml_for_hs.pdf (дата обращения 10.04.2020).

14. *Pangrazio L., Sefton-Green J.* The social utility of ‘data literacy.’ // Learning, Media and Technology, 2019. doi:10.1080/17439884.2020.1707223

15. *Schield M.* Information literacy, statistical literacy and data literacy. // IASSIST Quarterly, 28(2), 6–11, 2018.

16. *Srikant, S., & Aggarwal, V.* (2017). Introducing data science to school kids / Proceedings of the Conference on Integrating Technology into Computer Science Education, ITiCSE, 561–566. doi:10.1145/3017680.3017717

17. *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). C. 41–54. doi:10.1016/j.ijhcs.2019.03.006