

RU

## Направления исследований в области анализа образовательных данных в высшей школе: теоретический обзор

Семёнкина И. А., Прусакова П. В.

**Аннотация.** Данная публикация представляет собой обзор зарубежной англоязычной научно-педагогической литературы, цель которого – выявить актуальные направления исследований в области анализа образовательных данных в современной высшей школе. В обзоре рассмотрены факторы, обусловившие развитие анализа образовательных данных (далее – АОД) и аналитики обучения (далее – АО) в контексте процессов цифровой трансформации современного общества. Разбираются потенциал, проблемы и направления применения АОД и АО в высшем образовании в целом, а также в сфере анализа успеваемости и поведения обучающихся, усовершенствования образовательных программ, повышения эффективности системы высшего образования. Научная новизна обзора заключается в определении наиболее актуальных задач исследований АОД и выявлении перспективных направлений исследований в данной области всех субъектов образовательного процесса в высшей школе. В результате проанализированы работы 2017-2023 гг. по рассматриваемой тематике, описаны проблемы применения АОД, связанные с вопросами этики и конфиденциальности личных данных; актуальные методы АОД; опыт внедрения АОД в высшей школе.

EN

## Lines of research in the area of educational data mining in higher education: A theoretical review

Semyonkina I. A., Prusakova P. Y.

**Abstract.** The paper provides an overview of foreign English-language research literature on pedagogy, the aim of which is to identify the most relevant lines of research in the area of educational data mining in modern higher education. The review considers the factors that have caused the development of educational data mining (hereinafter EDM) and learning analytics (hereinafter LA) in the context of digital transformation processes in modern society. The paper discusses the potential, problems and directions of implementing EDM and LA in higher education in general, as well as in the field of academic performance and students' behavior, educational programs development and improving the education system efficiency. Scientific novelty of the review lies in identifying the most relevant tasks for EDM research and defining advanced research directions in this area for all actors of educational process in higher education. As a result, the authors analyzed research papers on the described subject area published in the period from 2017 to 2023 and described issues related to personal data ethics and privacy in the context of EDM implementation, the relevant methods of EDM, the experience of EDM implementation in higher education.

### Введение

Начало изучения образовательных данных в электронной форме связывают с развитием в США в 1980-х гг. интеллектуальных обучающих систем, позволяющих отслеживать действия обучающихся. В значительной степени анализ образовательных данных возник как анализ журналов взаимодействия обучающегося с компьютером.

В дальнейшем использование Интернета и программного обеспечения в образовании, создание государственных баз данных, содержащих информацию об обучающихся, электронных ресурсов привели к накоплению огромного количества разнообразных данных в образовательной сфере и многократному росту количества публикаций на темы анализа образовательных данных (АОД) в 2000-х гг. в таких изданиях, как *Journal of Learning Analytics*, *Computers and Education*, *Journal of Educational Data Mining*, *Journal of Artificial Intelligence in Education*.

Уже в 2000 г. вышло в свет исследование J. Beck, B. Woolf (2000), в котором было показано, как методы прогнозирования интеллектуального анализа данных в образовании можно использовать для разработки моделей обучающихся.

В 2001-2005 гг. исследователи предпринимали попытки применения АОД для изучения онлайн-курсов и разработки более эффективных систем электронного обучения (Zaiane, 2001; Tang, McCalla, 2005).

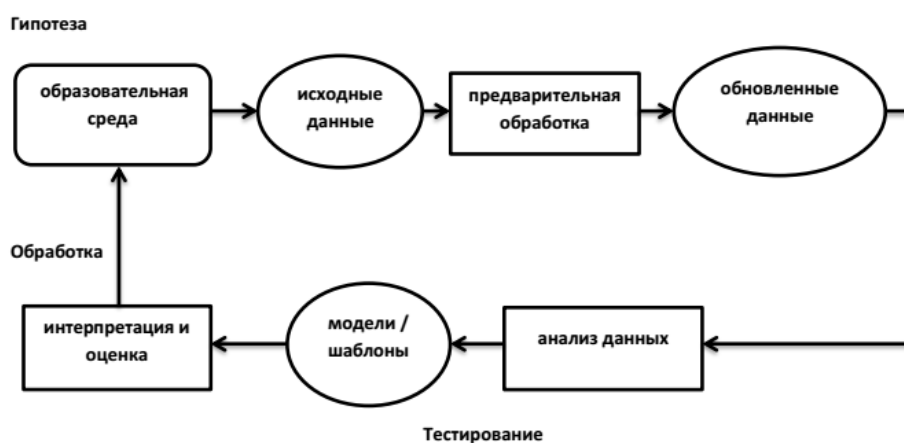
А. Merceron и К. Yacef (2003) в 2003 г. опубликовали работу, в которой представили инструменты АОД для повышения качества преподавания.

В 2004 г. была издана статья R. S. Baker, A. T. Corbett и К. R. Koedinger (2004) о применении методов прогнозирования в обучении, после которой стало очевидно, что анализ образовательных данных может давать конкретные результаты.

В 2007 г. вышла классическая обзорная статья “Educational Data Mining: A Survey from 1995 to 2005”, в которой были описаны цели анализа образовательных данных в условиях конкретного образовательного учреждения; типы доступных данных; методы анализа образовательных данных в зависимости от целей; перспективы применения анализа образовательных данных (Romero, Ventura, 2007).

В 2010 г. и 2020 г. С. Romero и S. Ventura (2010; 2020) опубликовали две статьи, рассмотрев в них подходы к АОД, учитывая современные реалии и новейшие технологии. По мнению авторов, новые типы образовательных сред, такие как смешанное обучение, виртуальные/расширенные среды, мобильное обучение, игровое обучение и т. д., также собирают огромное количество данных об обучающихся. Эта информация представляет собой важнейший источник образовательных данных, которые можно исследовать и использовать, чтобы понять закономерности процесса обучения.

Сложному процессу получения новых знаний об обучении посвящены работы О. Moscoso-Zea, А. Р. Sampedro, S. Luján-Mora (2016), Е. Alyahyan, D. Düşteğör (2020). По мнению авторов, АОД – это циклический процесс, состоящий из формулировки, проверки и уточнения гипотезы. Данный цикл получения новых знаний в учебных заведениях на основе сбора и обработки данных, генерируемых в образовательной среде, дальнейшей разработки модели знаний об образовательных явлениях и процессе обучения, интерпретации и внедрения модели представлен на Рисунке 1 (Alyahyan, Düşteğör, 2020).



**Рисунок 1.** Процесс получения новых знаний в учебных заведениях (переведено авторами. – И. С., П. П.)

Моделям знаний об образовательных явлениях и процессе обучения, которые разрабатываются на основе интеллектуальной обработки всего спектра образовательных данных, доступных учебным заведениям, посвящена работа (Rosé, McLaughlin, Liu et al., 2019).

Ряд исследователей (Baker, Inventado, 2014) выделяют в сфере образовательных данных два основных направления: анализ образовательных данных (Educational Data Mining, EDM) и аналитику обучения (Learning Analytics, LA).

Оба направления объединяет общий интерес к применению образовательных данных для развития теории обучения, усовершенствования организации материалов учебных курсов, управления образовательными ресурсами и оптимизации процесса проектирования образовательной среды. EDM можно определить как применение методов интеллектуального анализа данных (Data Mining, DM) к конкретному типу данных, получаемых в образовательной среде, для решения важных образовательных вопросов (Romero, Ventura, 2013). Аналитика обучения (АО) может быть определена как измерение, сбор, анализ и представление данных об обучающихся и их условиях обучения в целях понимания и оптимизации процесса обучения и среды, в которой оно происходит (Handbook of Learning Analytics, 2017). В это определение входят три важнейших элемента (Siemens, 2013): данные, анализ и действия.

Важно отметить, что, несмотря на различия между АОД и АО, эти два направления во многом совпадают как в целях исследователей, так и в используемых методах и приемах. Как междисциплинарные направления они включают поиск информации, анализ визуальных данных, интеллектуальный анализ данных, анализ социальных сетей, методы когнитивной психологии, психологии образования. АОД и АО можно представить как сферу пересечения трех направлений научных исследований: информатики, образования и статистики.

В последнее время исследования направлений применения АОД и АО в высшей школе главным образом были сфокусированы на определении целей и изучении сложности процесса анализа больших данных в образовании, тогда как влияние АОД на практику и теорию обучения оставалось ограниченным. Авторы видят

необходимость усилить влияние АОД и АО на практику обучения и перейти от исследовательских моделей образовательных явлений и процесса обучения к более целостным и интегративным исследованиям применения анализа образовательных данных в высшей школе на системном уровне.

Сегодня мы вынуждены констатировать, что возможности свободного использования инструментов АОД и АО заинтересованными специалистами ограничены множеством объективных факторов: недостаточность навыков анализа данных, этические, правовые и технические барьеры. Мы полагаем, что необходимо разрабатывать доступные инструменты АОД, а также культуру их применения для принятия обоснованных решений и повышения качества высшего образования.

Изучение образовательных данных активно проводится за рубежом, опубликованы результаты многочисленных исследований, предлагающих технические решения, описывающих практический опыт внедрения АОД для повышения основных показателей эффективности деятельности в науке и образовании (Attaran, Stark, Stottler, 2018; Sekli, De la Vega, 2021). В России АОД не получил такого широкого обсуждения, что обусловило выбор зарубежных публикаций для настоящего обзора.

Задача настоящего исследования – выявить актуальные направления исследований в области анализа образовательных данных в современной высшей школе за рубежом.

Материалом для анализа послужили свыше 40 печатных и электронных источников по анализу образовательных данных в системе высшего образования на английском языке с 2017 г. по 2023 г. включительно, посвященных потенциалу применения АОД в высшей школе, анализу поведения и прогнозированию успеваемости обучающихся, усовершенствованию образовательных программ и повышению качества обучения, этике и конфиденциальности личных данных, методам АОД, опыту внедрения АО за рубежом.

Авторами применялись следующие методы: изучение тематики научных публикаций по проблеме исследования, теоретический анализ, обобщение, систематизация.

Используемая систематизация направлений исследований в области анализа образовательных данных является в значительной степени относительной, так как в рамках проанализированных исследований многие тематические разделы пересекаются по смыслу. Каждый тематический блок мы выделяли, отталкиваясь от преобладающего содержания публикации и актуальности темы по оценкам авторов.

Остановимся подробнее на публикациях 2017-2023 гг., в которых рассмотрены направления исследований в области анализа образовательных данных в современной высшей школе. Направления исследований представлены в Таблице 1.

**Таблица 1.** Тематический обзор англоязычных научных публикаций по направлениям исследований в области анализа образовательных данных в высшей школе (2017-2023 гг.)

№	Тематика и направления исследований применения АОД в высшей школе	Источники
1.	Потенциал, проблемы и направления применения АОД в высшей школе	Daniel, 2019; Dawson, Poquet, Colvin et al., 2018; Hoel, Chen, 2018; Ifenthalter, Yau, 2020; Ang, Ge, Seng, 2020; Attaran, Stark, Stotler, 2018.
2.	Анализ успеваемости и поведения обучающихся с помощью АОД в высшей школе	Ahmad, Shahzadi, 2018; Azcona, Hsiao, Smeaton, 2019; Juhaňák, Zounek, Rohlíková, 2019; Musso, Hernández, Cascallar, 2020; Namoun, Alshanqiti, 2021; Shorfuzzaman, Hossain, Nazir et al., 2019; Waheed, Hassan, Aljohani et al., 2020; Zhang, Li, 2018.
3.	Анализ образовательных программ с помощью АОД в высшей школе	Brown, DeMonbrun, Teasley, 2018; Gottipati, Shankaraman, 2018; Heileman, Slim, Hickman et al., 2017; Hilliger, Aguirre, Miranda et al., 2020; Kitto, Sarathy, Gromov et al., 2020.
4.	Повышение эффективности высшей школы с помощью АОД	Attaran, Stark, Stotler, 2018; Dhankhar, Solanki, 2020; Hernández de Menéndez, Morales-Menendez, Escobar et al., 2022; Hooda, Rana, 2020; Tasmin, Muhammad, Aziati, 2020.
5.	Этика и конфиденциальность личных данных обучающихся в высшей школе при использовании АОД	Guan, Feng, Islam, 2023; West, Luzeckyj, Toohey et al., 2020; Kousa, Niemi, 2022; Kularski, Martin, 2021; Li, Chen, Sun et al., 2021; Solem, Dony, Herman et al., 2021; Mandinach, Jimerson, 2021; Tang, Zhang, 2020; Zeng, Li, Cao et al., 2020.
6.	Методы АОД в высшей школе	Ahuja, Jha, Maurya et al., 2018; Aleem, Gore, 2020; Alyahyan, Düşteğör, 2020; Bakhshinategh, Zaiane, ElAtia et al., 2018; Dimić, Predić, Rančić et al., 2017; Hooda, Rana, 2020; Vieira, Parsons, Byrd, 2018.
7.	Опыт внедрения АО в высшей школе за рубежом	Attaran, Stark, Stotler, 2018; Sonderlund, Hughes, Smith, 2018; Hooda, Rana, 2020; Hernández de Menéndez, Morales-Menendez, Escobar et al., 2022; Sekli, De la Vega, 2021; Tsai, Rates, Moreno-Marcos et al., 2020.

## Аналитическая часть

### 1. Потенциал, проблемы и направления применения АОД в высшей школе

Важнейший ресурс АОД для вузов – повышение эффективности системы высшего образования для всех участников образовательного процесса. Для обучающихся АОД может быть использован с целью регулирования процессов зачисления, информационной поддержки обучающихся, снижения рисков отчисления, выстраивания

индивидуальных образовательных траекторий, взаимодействия с выпускниками, управления финансовой помощью (Attaran, Stark, Stotler, 2018). Педагоги могут отслеживать вовлеченность обучающихся в учебный процесс, когнитивные, поведенческие и другие аспекты процесса обучения, академическую успеваемость обучающихся, получать обратную связь об эффективности конкретного курса.

АОД является основой развития доказательного менеджмента и доказательной педагогики, т. е. обеспечивает принятие эффективных решений администрациями вузов на основании интеллектуального анализа данных. В сфере образовательных данных D. Ifenthaler и J. Yau (2020) выделяют три профиля данных: профиль обучающегося (например, демографические данные обучающихся), профиль обучения (например, успеваемость обучающихся в системе управления обучением) и профиль учебной программы (например, результаты обучения по курсу и последовательность курсов в рамках учебного плана).

Перспективы применения АОД в высшей школе выглядят многообещающими, однако исследователи указывают на социальные, технические и организационные проблемы практического внедрения анализа данных. Технические проблемы обусловлены прежде всего тем, что для крупных вузов необходим доступ к высокопроизводительной вычислительной инфраструктуре, способной обрабатывать большие объемы данных для сбора, хранения, обработки и визуализации (Ang, Ge, Seng, 2020). Кроме того, существует проблема функциональной совместимости систем данных в рамках образовательных учреждений и различных форм хранения данных в разрозненных базах данных (Daniel, 2019). Социальные проблемы могут быть вызваны предвзятым отношением к оценке способностей обучающегося на основе АОД и лишением обучающегося доступа к определенным образовательным программам в автоматическом режиме с негативными последствиями для профессиональной карьеры (Hoel, Chen, 2018). Кроме того, руководители университетов должны понимать потенциал и преимущества внедрения комплексной системы АОД, последовательно выстраивать грамотную политику в данном направлении (Attaran, Stark, Stotler, 2018). Организационные проблемы практического внедрения анализа данных объясняются во многом осторожностью образовательных организаций. Не исключена вероятность опубликования нежелательных данных, например низкой результативности обучения по сравнению с ведущими университетами страны или даже средними показателями по стране (Dawson, Poquet, Colvin et al., 2018).

## **2. Анализ успеваемости и поведения обучающихся с помощью АОД в высшей школе**

Учебные заведения имеют доступ к различным данным об обучающихся, начиная от демографических и заканчивая отчетами о работе в системах управления обучением, системами интеллектуального обучения, массовыми открытыми онлайн-курсами, информационными системами для обучающихся. Объединение всех имеющихся данных позволяет получить цифровой след обучающегося, анализ которого помогает университетам лучше понимать поведение студентов, раскрывать их академический потенциал, прогнозировать успеваемость (Azcona, Hsiao, Smeaton, 2019).

Своевременное прогнозирование успеваемости позволяет выявлять обучающихся с низкой успеваемостью, тем самым давая педагогам возможность оказывать влияние на процесс обучения на раннем этапе и принимать необходимые меры (Zhang, Li, 2018; Shorfuzzaman, Hossain, Nazir et al., 2019).

Анализ данных об успеваемости тесно связан с концепцией образования, основанного на результатах. Эта популярная сегодня модель обучения является студентоцентрированной и сфокусирована на оценке результатов освоения программы путем определения уровня знаний, навыков и поведения, которые выпускник должен реализовать по завершении программы.

Исследователи обращают внимание на тот факт, что успеваемость остается серьезной проблемой в высшем образовании, в особенности учитывая низкие баллы обучающихся и растущий процент отчислений даже в ведущих мировых университетах (Juhaňák, Zounek, Rohlíková, 2019; Namoun, Alshantiti, 2021; Zhang, Li, 2018).

Разработано множество моделей достаточно точного прогнозирования успеваемости, используя различные динамические и статические характеристики обучающихся (стратегии обучения, восприятие социальной поддержки, мотивация, социально-демографические факторы, состояние здоровья, показатели академической успеваемости, средний балл аттестата и пр.); нейронные сети и машинное обучение для анализа действий обучающихся, связанных с их навигацией по LMS (Ahmad, Shahzadi, 2018; Musso, Hernández, Cascallar, 2020; Waheed, Hassan, Aljohani et al., 2020).

Таким образом, применение АОД для прогнозирования успеваемости обучающихся может способствовать увеличению соотношения числа окончивших учебное заведение к числу поступивших.

## **3. Анализ образовательных программ с помощью АО в высшей школе**

Анализ образовательных программ возник как отдельное направление аналитики обучения, специализирующееся на сборе и анализе образовательных данных для принятия обоснованных решений по развитию учебных программ и повышению качества этих программ (Hilliger, Aguirre, Miranda et al., 2022). К такого рода данным относятся структура программы, последовательность презентации контента, учебные ресурсы, виды учебной деятельности, организация оценивания по курсу (Gottipati, Shankaraman, 2018; Brown, DeMonbrun, Teasley, 2018). Эти образовательные данные могут быть получены из нескольких источников информации на уровне курса и программы. Данные на уровне курса состоят из учебных программ по дисциплинам, оценок обучающихся, тогда как данные уровня программы состоят из выбора курса, карт учебного плана и матриц результатов обучающихся. Эти матрицы описывают взаимосвязь между основными курсами и компетенциями, которые обучающиеся должны развивать в соответствии с профилем программы (Hilliger, Aguirre, Miranda et al., 2020).

В аспекте приобретения компетенций, предусмотренных программой каждого конкретного курса, ряд исследователей сосредоточили свои усилия на составлении карт развития навыков в процессе освоения учебных программ, анализе описаний курсов с точки зрения таксономии навыков и уровней прогресса (Kitto, Sarathy, Gromov et al., 2020; Gottipati, Shankararaman, 2018).

Разработаны различные типы инструментов анализа образовательных программ для конкретных профилей подготовки, преподавателей и обучающихся, направленные на улучшение учебных программ и результатов обучения. Например, некоторые инструменты позволяют анализировать данные на уровне программы, включая соотношение числа зачисленных на курс и закончивших его; время, затраченное на завершение курса; предоставляют интерактивную визуализацию потоков обучающихся в рамках академических программ, графическое изображение алгоритма прохождения курса (Heileman, Slim, Hickman et al., 2017).

Несмотря на определенные успехи в области применения АО, в настоящее время необходимо проведение дальнейших исследований, основанных на проектировании, для разработки и оценки соответствующих инструментов непрерывного усовершенствования курсов и программ.

#### ***4. Повышение эффективности высшей школы с помощью АОД***

Принятие обоснованных решений на основе анализа больших данных стало одной из важнейших задач в сфере высшего образования в новых реалиях. Традиционно решения, принимаемые администрацией и преподавательским составом высших учебных заведений, опирались на предположения и определенные гипотезы. Этот процесс требует значительных затрат времени и ресурсов, а также ограничен компетенциями, управленческими и другими качествами основных акторов. С помощью ресурсов аналитики специалисты создают новые схемы и технологии, которые помогают вузам интегрировать имеющиеся данные в единую систему и получить исчерпывающую информацию о работе организации как единого организма (Hooda, Rana, 2020).

В специальной литературе также встречается термин академическая аналитика (Academic Analytics, AA), в которой аналитические методы применяются прежде всего для принятия решений, отвечающих институциональным, функциональным, административным и финансовым потребностям вузов (Dhankhar, Solanki, 2020).

В сфере использования AA для повышения эффективности высших учебных заведений M. Attaran, J. Stark, D. Stotler (2018) выделяют несколько перспективных аспектов: управление зачислением обучающихся на курс, управление финансовой помощью, управление успеваемостью и оценивание обучающихся с использованием цифровых технологий, умный класс, анализ обучения, расписание занятий, продвижение университета.

Среди направлений применения академической аналитики также можно рассматривать сохранение и развитие профессорско-преподавательского состава высшей школы, пересмотр программ обучения, рационализация деятельности подразделений учебных заведений, стратегическое планирование, вопросы финансирования и получения грантов, контакты с выпускниками, участие в конкурсах, вопросы сотрудничества, повышение квалификации сотрудников (Hernández de Menéndez, Morales-Menendez, Escobar et al., 2022; Dhankhar, Solanki, 2020; Tasmin, Muhammad, Aziati, 2020).

Учитывая запрос общества на повышение качества высшего образования, нет сомнений в том, что применение академической аналитики может положительно влиять на деятельность вузов, преобразовывать и значительно усовершенствовать их, особенно в отношении стратегического планирования и политики в системе образования.

#### ***5. Этика и конфиденциальность личных данных обучающихся в высшей школе при использовании АОД***

Этика образовательных данных относится к разделу этики, определяющему принципы сбора, управления, обмена и использования образовательных данных. Соблюдение этих принципов должно предотвращать негативное влияние анализа образовательных данных как на общество в целом, так и на его отдельных членов. Этика образовательных данных может обеспечить рациональное использование данных, правильный анализ и интерпретацию выводов (Mandinach, Jimerson, 2021; Kousa, Niemi, 2022).

X. Guan, X. Feng, A. Islam (2023) выделяют три основные проблемы в этике образовательных данных: (а) нарушение конфиденциальности при сборе, хранении и обмене данными, (б) лишение образовательных субъектов возможности делать независимый выбор путем прогнозирования образовательных данных, (в) применение образовательных данных «навязывает» оценку обучающихся исключительно по жестким критериям данных, отсутствует «способность к забыванию» оценки обучающегося внутри базы образовательных данных, у образовательных субъектов могут формироваться о себе ложные стереотипы. Исследователи обращают внимание на такие потенциальные риски использования АОД не по назначению, как, например, продажа данных, так называемая социальная сортировка обучающихся на группы по признакам уровня доходов, расы, этнической принадлежности, социального статуса, усиление и распространение кибернасилия (West, Luzeckij, Toohey et al., 2020).

С. М. Kularski, F. Martin (2021) акцентируют уязвимость конфиденциальной информации в образовательных данных. Хранение данных без их эффективность защиты может привести к намеренной или случайной утечке информации, в дальнейшем к ее использованию против обучающихся или представителей образовательного учреждения и, как следствие, создает потенциальную угрозу для них.

По мнению А. Magdy, С. Dony (2019), если преподаватели чрезмерно полагаются на функцию прогнозирования, обеспечиваемую образовательными данными, понижается уровень их критического мышления, что в конечном итоге может деструктивно повлиять на профессиональное развитие педагогов.

Специалисты выражают обеспокоенность тем, что, несмотря на повышение эффективности обучения и точность прогнозов на основе АОД, значительно ослабляется возможность самостоятельного мышления обучающихся и инноваций методом проб и ошибок во время обучения. В результате сужается кругозор обучающихся, создается информационная капсула, что, в свою очередь, может негативно сказаться на воспитании у молодежи нравственных ценностей и формировании мировоззрения (Li, Chen, Sun et al., 2021).

Важно осознавать, что обучающиеся в большинстве случаев являются молодыми людьми, личность которых продолжает формироваться во время обучения, поэтому их автоматическое распределение на категории на основе образовательных данных крайне опасно. Более того, длительное хранение образовательных данных об обучающихся может способствовать закреплению стереотипов о способностях обучающихся и снижать мотивацию к развитию и самосовершенствованию (Tang, Zhang, 2020).

Авторы статьи разделяют подход к данной проблеме авторов работы (Zeng, Li, Cao et al., 2020), которые предполагают, что создание центров управления данными в учебных учреждениях, использование распределенного хранения в блокчейне, шифрование с секретным ключом и другие методы и технологии могут защитить конфиденциальность данных педагогов и обучающихся.

Несовершенство нормативной базы в отношении защиты персональных данных обучающихся и сотрудников вузов является веским аргументом противников внедрения АОД в высшей школе.

### **6. Методы АОД в высшей школе**

Термин «методы АОД» относится к ряду специальных аналитических методов, которые применяются исключительно к образовательным данным.

М. Hooda, С. Rana (2020) в своей работе определяют 14 методов анализа образовательных данных и аналитики обучения, используемых для прогнозирования, распределения по категориям, мониторинга деятельности студентов и преподавателей: классификация, кластеризация, правила ассоциации, корреляция, статистика, анализ текста, геймификация, анализ социальных сетей, визуализация, отбор данных для изучения, моделирование, регрессия, изучение моделей последовательности действий, выявление аномалий.

А. Aleem, М. М. Gore (2020) выделяют шесть основных категорий методов АОД: 1) отбор данных для изучения; 2) методы прогнозирования; 3) методы анализа взаимосвязей; 4) методы обнаружения структуры; 5) моделирование; 6) другие методы (анализ социальных сетей, анализ текстовых данных и др.).

**Отбор данных для изучения** предполагает выделение («подсвечивание») полезной информации с помощью предоставления данных, например используя обобщение, визуализацию и интерактивные интерфейсы (Vieira, Parsons, Byrd, 2018).

**Методы прогнозирования** являются наиболее востребованными методами для решения таких проблем, как отчисление обучающихся или низкие показатели их успеваемости (Alyahyan, Düşteğör, 2020).

**Методы анализа взаимосвязей** используются для выявления отношений между характерными признаками явлений и процессов в базах образовательных данных. Также обнаруживаются значимые взаимосвязи между факторами, которые могут быть полезны для улучшения методологии обучения или образовательной системы (Dimić, Predić, Rančić et al., 2017).

**Методы обнаружения структуры** используют механизм неконтролируемой классификации для поиска структуры в наборе данных. Методы этой категории являются вторыми по частоте использования методами АОД после методов прогнозирования (Ahuja, Jha, Maurya et al., 2018).

**Метод моделирования** (создания модели обучающегося) представляет процесс воспроизведения когнитивной деятельности обучающегося, включая ошибки и неверное понимание материала курса, стили обучения, аффективные, когнитивные и метакогнитивные факторы. Полученная модель может использоваться для прогнозирования успеваемости обучающихся, достижения результатов обучения, адаптации трудоемкости курса (Bakhshinategh, Zaiane, ElAtia et al., 2018).

Несомненно, применение разнообразных методов АОД и АО предполагает соответствие программного обеспечения поставленным задачам в вузе для сбора и анализа данных, а также компетентность сотрудников.

### **7. Опыт внедрения АО в высшей школе за рубежом**

Опыт внедрения АО описан в многочисленных зарубежных публикациях с разбором конкретных ситуаций (кейсов) из практики высшей школы.

Так, М. Attaran, J. Stark, D. Stotler (2018) рассматривают примеры из практики использования АО в крупнейших университетах США (Университет штата Аризона, Университет штата Мичиган, Университет штата Джорджия, Университет Джона Хопкинса, Университет штата Делавэр, Калифорнийский университет и др.) в зависимости от целей ее применения на этапах взаимодействия университета с обучающимися: на этапе приема в университет (анализ вероятности принятия абитуриентами решения о поступлении, поиск способов мотивации к поступлению и др.); на этапе обучения (поиск способов увеличения процента выпускников и количества перешедших на второй курс, сокращения времени обучения до получения диплома; снижения разницы в успеваемости различных групп обучающихся и др.), на этапе развития вуза (поиск способов повышения качества работы персонала вуза, экономии на оплате труда, обеспечения роста ежегодных доходов от спонсоров и др.).

А. Sonderlund, Е. Hughes, J. R. Smith (2018) исследуют применение аналитических систем в вузах зарубежных стран (США, Бразилия, Тайвань, Южная Корея) с целью оценки эффективности активного вмешательства

преподавателей и систем аналитики обучения в учебный процесс. Рассматриваются такие модели вмешательства, как оказание преподавателями консультационной помощи обучающимся группы риска, автоматическая отправка обучающимся личных сообщений с основной информацией о ходе обучения, разработка электронных сред, позволяющих обучающимся самостоятельно оценивать свою образовательную активность и сравнивать свои показатели с показателями других обучающихся по этой же программе; автоматическое включение обучающихся в группы риска с последующей отправкой им информации о принадлежности к этим группам по телефону и электронной почте; анализ вовлеченности обучающегося в учебный процесс и др.

Y. Tsai и соавторы (Tsai, Rates, Moreno-Marcos et al., 2020) в своей работе приводят результаты опроса сотрудников 249 вузов из 38 стран Европы по поводу эффективности внедрения аналитики обучения с точки зрения ее влияния на качество решения текущих проблем, возможность получения информации для принятия организационных решений, выводы о необходимости внесения изменений в образовательные программы, изменение восприятия учебного процесса студентами. Также данная работа исследует факторы, влияющие на успешность внедрения аналитики обучения: принятие руководством вуза и преподавательским составом изменений организации учебного процесса как результата внедрения АО; законодательная основа для внедрения аналитических систем; наличие средств сохранения конфиденциальности; соблюдение этических принципов; размер инвестиций, уровень технологических возможностей вузов.

Согласно исследованию M. Sekli и I. De la Vega (2021), на успешность внедрения систем анализа больших данных в высшей школе влияют совместимость аналитических систем с другими системами учреждения, хорошо структурированная среда данных, возможность привлечения сторонних специалистов и ресурсов. В то время как поддержка руководства вуза, уровень сложности аналитической системы и наличие конкуренции не влияют на успешность внедрения систем анализа.

Согласно исследованиям, наиболее частыми целями применения аналитики обучения являются следующие: сокращение процента отчисленных, прогнозирование и улучшение показателей академической успеваемости, персонализация программ обучения, анализ поведения обучающихся, поддержка обучающихся группы риска путем вмешательства в учебный процесс, оценка поведенческих стереотипов обучающихся (Hooda, Rana, 2020; Hernández de Menéndez, Morales-Menendez, Escobar et al., 2022).

Таким образом, проанализированные работы исследователей свидетельствуют о наличии успешного опыта внедрения АО в зарубежных вузах. Были зафиксированы следующие положительные результаты применения АО: повышение количества обучающихся, поступивших в вуз; уменьшение процента отчисленных и, соответственно, увеличение процента выпускников; улучшение показателей академической успеваемости; сокращение длительности обучения (времени от поступления до окончания курса); более легкий процесс принятия организационных решений по работе вуза; увеличение доходов вузов от спонсоров.

## **Заключение**

Проведенный анализ литературы показывает значительный интерес исследователей к проблематике применения АОД и АО в современной высшей школе за рубежом. Среди наиболее актуальных направлений исследований следует выделить: потенциал и проблемы применения АОД в высшей школе, анализ успеваемости и поведения обучающихся, анализ образовательных программ, повышение эффективности высшей школы с помощью АОД, этика и конфиденциальность персональных данных обучающихся, методы АОД, практический опыт внедрения АО.

Огромный массив образовательных данных, имеющихся в распоряжении высших учебных заведений, предоставляет широкие возможности для развития системы образования в целом, включая усовершенствование учебного процесса, управления, финансирования, планирования деятельности вузов. Понимание закономерностей сложнейших процессов, реализуемых учебными заведениями, перспективных тенденций развития института образования, выявление эффективных паттернов и нерезультативных практик с помощью анализа образовательных данных может позволить вузам принимать обоснованные, взвешенные решения, значительно сократив затраты времени и других ресурсов.

Исследования в области анализа образовательных данных также способны дать новый импульс развитию педагогики и психологии высшей школы с учетом возможностей инновационных технологий, позволяющих в реальном времени исследовать работу человеческого мозга (например, нейрокомпьютерные интерфейсы).

Практический опыт внедрения АОД ведущими зарубежными университетами показывает, что интеллектуальный анализ образовательных данных может оказывать положительное влияние на прогнозирование академической успеваемости обучающихся, снижение количества отчисленных, выявление трудностей, возникающих у обучающихся, и способствует раннему вмешательству в ситуации риска. АОД позволяет развивать персонализацию обучения на основе достоверных данных профиля обучающихся, включающих комплекс факторов, динамические и статические характеристики.

Проблемы эффективного внедрения анализа образовательных данных связаны со сложностью формирования конкретных целей; правовым обеспечением вопросов защиты личных данных участников образовательного процесса; риском формирования предвзятого отношения к обучающимся на основе оценок способностей по результатам АОД; трудоемкостью функционирования информационно-технологической инфраструктуры в вузе, отсутствием подготовленных специалистов в штате организаций.

Приоритетными направлениями дальнейших исследований в этой связи представляются:

- систематизация и стандартизация образовательных данных, имеющихся в доступе высших учебных заведений;
- нормативно-правовое обеспечение применения анализа образовательных данных в высшей школе;
- развитие отечественных технологий анализа образовательных данных в высшей школе;
- развитие теории обучения и психологии высшей школы на основе анализа образовательных данных и аналитики обучения;
- оптимизация проектирования образовательной среды на основе анализа образовательных данных и аналитики обучения.

## Источники | References

1. Ahmad Z., Shahzadi E. Prediction of Students' Academic Performance Using Artificial Neural Network // *Bulletin of Education*. 2018. Vol. 40. Iss. 3.
2. Ahuja R., Jha A., Maurya R., Srivastava R. Analysis of Educational Data Mining // *Advances in Intelligent Systems and Computing*. 2018. Vol. 741.
3. Aleem A., Gore M. M. Educational Data Mining Methods: A Survey // 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT) (Gwalior, India, 2020). 2020. DOI: 10.1109/CSNT48778.2020.9115734
4. Alyahyan E., Düşteğör D. Predicting Academic Success in Higher Education: Literature Review and Best Practices // *International Journal of Educational Technology in Higher Education*. 2020. Vol. 17. Iss. 3.
5. Ang L., Ge F., Seng K. Big Educational Data & Analytics: Survey, Architecture and Challenges // *IEEE Access*. 2020. Vol. 8.
6. Attaran M., Stark J., Stottler D. Opportunities and Challenges for Big Data Analytics in American Higher Education – A Conceptual Model for Implementation // *Industry and Higher Education*. 2018. Vol. 32. Iss. 3.
7. Azcona D., Hsiao I.-H., Smeaton A. F. Detecting Students-at-Risk in Computer Programming Classes with Learning Analytics from Students' Digital Footprints // *User Modeling and User-Adapted Interaction*. 2019. Vol. 29. Iss. 2.
8. Baker R. S., Corbett A. T., Koedinger K. R. Detecting Student Misuse of Intelligent Tutoring Systems // *Proceedings of the 7th International Conference on Intelligent Tutoring Systems (Maceio, Alagoas, Brazil, August 30 – September 3, 2004)*. 2004. DOI: 10.1007/978-3-540-30139-4\_50
9. Baker R. S., Inventado P. S. Educational Data Mining and Learning Analytics // *Learning Analytics* / ed. by J. Larusson, B. White. N. Y., 2014.
10. Bakhshinategh B., Zaiane O. R., ElAtia S., Ipperciel D. Educational Data Mining Applications and Tasks: A Survey of the Last 10 Years // *Education and Information Technologies*. 2018. Vol. 23. Iss. 1.
11. Beck J., Woolf B. High-Level Student Modeling with Machine Learning // *Proceedings of the 5th International Conference of Intelligent Tutoring Systems (Germany, Berlin, June 19-23, 2000)*. 2000. DOI: 10.1007/3-540-45108-0\_62
12. Brown M., DeMonbrun R. M., Teasley S. Taken Together: Conceptualizing Students' Concurrent Course Enrollment Across the Post-Secondary Curriculum Using Temporal Analytics // *Journal of Learning Analytics*. 2018. Vol. 5. Iss. 3.
13. Daniel B. K. Big Data and Data Science: A Critical Review of Issues for Educational Research // *British Journal of Educational Technology*. 2019. Vol. 50. Iss. 1.
14. Dawson S., Poquet O., Colvin C., Rogers T., Pardo A., Gasevic D. Rethinking Learning Analytics Adoption through Complexity Leadership Theory // *LAK'18 Proceedings of the 8th International Conference on Learning Analytics and Knowledge (Australia, Sydney, March, 2018)*. 2018. DOI: 10.1145/3170358.3170375
15. Dhankhar A., Solanki K. State of the Art of Learning Analytics in Higher Education // *International Journal of Emerging Trends in Engineering Research*. 2020. Vol. 8. Iss. 3.
16. Dimić G., Predić B., Rančić D., Petrović V., Maček N., Spalević P. Association Analysis of Moodle E-Tests in Blended Learning Educational Environment // *Computer Applications in Engineering Education*. 2017. Vol. 26. Iss. 3.
17. Gottipati S., Shankaraman V. Competency Analytics Tool: Analyzing Curriculum Using Course Competencies // *Education and Information Technologies*. 2018. Vol. 23. Iss. 10.
18. Guan X., Feng X., Islam A. The Dilemma and Countermeasures of Educational Data Ethics in the Age of Intelligence // *Humanities and Social Sciences Communications*. 2023. Vol. 10.
19. *Handbook of Learning Analytics* / ed. by C. Lang, G. Siemens, A. Wise, D. Gašević. 1st ed. Society for Learning Analytics Research (soLAR), 2017. DOI: 10.18608/hla17
20. Heileman G. L., Slim A., Hickman M., Abdallah C. T. Characterizing the Complexity of Curricular Patterns in Engineering Programs // *ASEE Annual Conference & Exposition (Columbus, Ohio, June 24-28, 2017)*. 2017. DOI: 10.18260/1-2--28029
21. Hernández de Menéndez M., Morales-Menendez R., Escobar C., Ramirez-Mendoza R. Learning Analytics: State of the Art // *International Journal on Interactive Design and Manufacturing (IJIDeM)*. 2022. Vol. 16. Iss. 2.
22. Hilliger I., Aguirre C., Miranda C., Celis S., Pérez-Sanagustín M. Lessons Learned from Designing a Curriculum Analytics Tool for Improving Student Learning and Program Quality // *Journal of Computing in Higher Education*. 2022. Vol. 34.
23. Hilliger I., Aguirre C., Miranda C., Celis S., Pérez-Sanagustín M. Design of a Curriculum Analytics Tool to Support Continuous Improvement Processes in Higher Education // *LAK'20: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (Frankfurt, Germany, March 23-27, 2020)*. 2020. DOI: 10.1145/3375462.3375489



24. Hoel T., Chen W. Privacy and Data Protection in Learning Analytics Should Be Motivated by an Educational Maxim – towards a Proposal // *Research and Practice in Technology Enhanced Learning*. 2018. Vol. 13. Iss. 20.
25. Hooda M., Rana C. Learning Analytics Lens: Improving Quality of Higher Education // *International Journal of Emerging Trends in Engineering Research*. 2020. Vol. 8. Iss. 5.
26. Ifenthalter D., Yau J. Reflections on Different Learning Analytics Indicators for Supporting Study Success // *International Journal of Learning Analytics and Artificial Intelligence for Education*. 2020. Vol. 2. Iss. 2.
27. Juhaňák L., Zounek J., Rohlíková L. Using Process Mining to Analyze Students' Quiz-Taking Behavior Patterns in a Learning Management System // *Computers in Human Behavior*. 2019. Vol. 92.
28. Kitto K., Sarathy N., Gromov A., Liu M., Musial K., Shum S. B. Towards Skills-Based Curriculum Analytics: Can We Automate the Recognition of Prior Learning? // *LAK'20: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (Frankfurt, Germany, March 23-27, 2020)*. 2020. DOI: 10.1145/3375462.3375526
29. Kousa P., Niemi H. AI Ethics and Learning: EdTech Companies' Challenges and Solutions // *Interactive Learning Environments*. March 1, 2022. DOI: 10.1080/10494820.2022.2043908
30. Kularski C. M., Martin F. Online Student Privacy in Higher Education: A Systematic Review of the Research // *American Journal of Distance Education*. 2021. Vol. 36. Iss. 3.
31. Li Y., Chen X., Sun D., Zhu Y., Zhai X. From "Transparent People" to "Practitioner": Challenges and Responses to Information Security in Higher Education: Implications from 2021 EDUCAUSE Horizon Report // *Journal of Distance Education*. 2021. Vol. 3.
32. Magdy A., Dony C. 1st ACM SIGSPATIAL Workshop on Geo Computational Thinking in Education. 2019. URL: [https://www.sigspatial.org/wp-content/uploads/special-issues/11/3/Report02\\_GeoEd.pdf](https://www.sigspatial.org/wp-content/uploads/special-issues/11/3/Report02_GeoEd.pdf)
33. Mandinach E., Jimerson J. Data Ethics in Education: A Theoretical, Practical, and Policy Issue // *Studia Paedagogica*. 2021. Vol. 26. Iss. 4.
34. Merceron A., Yacef K. A Web-Based Tutoring Tool with Mining Facilities to Improve Learning and Teaching // *11th International Conference on Artificial Intelligence in Education / ed. by F. Verdejo and U. Hoppe*. Sydney: IOS Press, 2003.
35. Moscoso-Zea O., Sampedro A. P., Luján-Mora S. Datawarehouse Design for Educational Data Mining // *2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET) (Turkey, Istanbul, 8-10 September, 2016)*. 2016. DOI: 10.1109/ITHET.2016.7760754
36. Musso M. F., Hernández C. F. R., Cascallar E. C. Predicting Key Educational Outcomes in Academic Trajectories: A Machine-Learning Approach // *Higher Education*. 2020. Vol. 80.
37. Namoun A., Alshantqi A. Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review // *Applied Sciences*. 2021. Vol. 11. Iss. 1.
38. Romero C., Ventura S. Data Mining in Education // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2013. Vol. 3. Iss. 1.
39. Romero C., Ventura S. Educational Data Mining: A Review of the State of the Art // *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2010. Vol. 40. Iss. 10.
40. Romero C., Ventura S. Educational Data Mining: A Survey from 1995 to 2005 // *Expert Systems with Applications*. 2007. Vol. 33. Iss. 1.
41. Romero C., Ventura S. Educational Data Mining and Learning Analytics: An Updated Survey // *WIREs Data Mining and Knowledge Discovery*. 2020. Vol. 10. Iss. 3.
42. Rosé C. P., McLaughlin E. A., Liu R., Koedinger K. R. Explanatory Learner Models: Why Machine Learning (Alone) Is Not the Answer // *British Journal of Educational Technology*. 2019. Vol. 50. Iss. 6.
43. Sekli M., De la Vega I. Adoption of Big Data Analytics and Its Impact on Organizational Performance in Higher Education Mediated by Knowledge Management // *Journal of Open Innovation: Technology, Market and Complexity*. 2021. Vol. 7. Iss. 4.
44. Shorfuzzaman M., Hossain M. S., Nazir A., Muhammad G., Alamri A. Harnessing the Power of Big Data Analytics in the Cloud to Support Learning Analytics in Mobile Learning Environment // *Computers in Human Behaviour*. 2019. Vol. 92.
45. Siemens G. Learning Analytics: The Emergence of a Discipline // *American Behavioral Scientist*. 2013. Vol. 57. Iss. 10.
46. Solem M., Dony C., Herman T., Leon K., Magdy A., Nara A., Ray W., Rey S., Russel R. Building Educational Capacity for Inclusive Geocomputation: A Research-Practice Partnership in Southern California // *Journal of Geography*. 2021. Vol. 120. Iss. 4.
47. Sonderlund A., Hughes E., Smith J. R. The Efficacy of Learning Analytics Interventions in Higher Education – a Systematic Review // *British Journal of Educational Technology*. 2018. Vol. 50.
48. Tang H., Zhang J. Limits of Big Data Application in Education // *Journal of East China Normal University (Educational Sciences)*. 2020. Vol. 38. Iss. 10.
49. Tang T., McCalla G. Smart Recommendation for an Evolving E-Learning System: Architecture and Experiment // *International Journal on E-Learning*. 2005. Vol. 4. Iss. 1.
50. Tasmin R., Muhammad N. R., Aziati A. Big Data Analytics Applicability in Higher Learning Educational System // *IOP Conference Series: Materials Science and Engineering (Kota Bharu, Kelantan, 2020)*. 2020. DOI: 10.1088/1757-899X/917/1/012064
51. Tsai Y., Rates D., Moreno-Marcos P., Merino P., Jivet I., Scheffel M., Drachsler H., Delago-Kloos C., Gasevic D. Learning Analytics in European Higher Education – Trends and Barriers // *Computers & Education*. 2020. Vol. 155.

52. Vieira C., Parsons P., Byrd V. Visual Learning Analytics of Educational Data: A Systematic Literature Review and Research Agenda // Computers & Education. 2018. Vol. 122.
53. Waheed H., Hassan S. U., Aljohani N. R., Hardman J., Alelyani S., Nawaz R. Predicting Academic Performance of Students from VLE Big Data Using Deep Learning Models // Computers in Human Behavior. 2020. Vol. 104.
54. West D., Luzeckyj A., Toohey D., Vanderlelie J., Searle B. Do Academics and University Administrators Really Know Better? The Ethics of Positioning Student Perspectives in Learning Analytics // Australasian Journal of Educational Technology. 2020. Vol. 36. Iss. 2.
55. Zaiane O. Web Usage Mining for a Better Web-Based Learning Environment. 2001. URL: <https://webdocs.cs.ualberta.ca/~zaiane/postscript/CATE2001.pdf>
56. Zeng Z., Li Y., Cao Y., Zhao Y., Zhong J., Sidorov D., Zeng X. Blockchain Technology for Information Security of the Energy Internet: Fundamentals, Features, Strategy and Application // Energies. 2020. Vol. 13. Iss. 4.
57. Zhang L., Li K. F. Education Analytics: Challenges and Approaches // 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA) (Krakow, Poland, 2018). 2018. DOI: 10.1109/WAINA.2018.00086

### Информация об авторах | Author information



**Семёнкина Ирина Артуровна<sup>1</sup>**, к. психол. н., доц.  
**Прусакова Полина Валентиновна<sup>2</sup>**  
<sup>1,2</sup> Московский политехнический университет



**Semyonkina Irina Arturovna<sup>1</sup>**, PhD  
**Prusakova Polina Valentinovna<sup>2</sup>**  
<sup>1,2</sup> Moscow Polytechnic University

<sup>1</sup> [isemyonkina@mail.ru](mailto:isemyonkina@mail.ru), <sup>2</sup> [pvk1985@gmail.com](mailto:pvk1985@gmail.com)

### Информация о статье | About this article

Дата поступления рукописи (received): 12.05.2023; опубликовано (published): 13.07.2023.

**Ключевые слова (keywords):** анализ образовательных данных; аналитика обучения; высшее образование; анализ успеваемости и поведения обучающихся; субъекты образовательного процесса; этика и конфиденциальность личных данных; educational data mining; learning analytics; higher education; analytics in the field of academic performance and students' behavior; actors of educational process; personal data ethics and privacy.