

Главацкий С. Т., Бурькин И. Г.

**ПРЕПОДАВАНИЕ ФУНДАМЕНТАЛЬНЫХ ОСНОВ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА КАК РЕАЛИЗАЦИЯ
КОНЦЕПЦИИ НОВОГО НАУЧНОГО ЗНАНИЯ**

Сергей Тимофеевич Главацкий

кандидат физико-математических наук, доцент

glavatsky_st@mail.ru

ФГБОУ ВО «Московский государственный университет имени

М. В. Ломоносова», Россия, Москва

Илья Геннадиевич Бурькин

научный сотрудник

Iliia.Burykin@sdo.msu.ru

ФГБОУ ВО «Московский государственный университет имени

М. В. Ломоносова», Россия, Москва

**TEACHING THE FUNDAMENTAL BASES OF ARTIFICIAL
INTELLIGENCE AS THE IMPLEMENTATION OF THE CONCEPT OF
NEW SCIENTIFIC KNOWLEDGE**

Sergei Timofeevich Glavatsky

M. V. Lomonosov Moscow State University, Russia, Moscow

Iliia Gennadievich Burykin

M. V. Lomonosov Moscow State University, Russia, Moscow

Аннотация. В статье рассматривается современная ситуация развития общества, когда оно вплотную подошло к восприятию методов и средств искусственного интеллекта (ИИ) в повседневной жизни как к реализации принципов нового научного знания (учения), базирующегося на фундаменте «Науки о данных». Предложена авторская концепция построения образовательных траекторий при подготовке специалистов в сфере ИИ.

***Abstract.** The article examines the current situation of the development of society, when it came close to the perception of methods and means of artificial intelligence (AI) in everyday life as the implementation of the principles of new scientific knowledge (teaching), based on the foundation of "Data Science". The author's concept of constructing educational trajectories for training specialists in the field of AI is offered.*

***Ключевые слова:** искусственный интеллект, высшее образование, ученый по данным, инженер по данным, аналитик данных.*

***Keywords:** Artificial Intelligence, Higher Education, Data Scientist, Data Engineer, Data Analyst.*

Сегодня человеческое общество достигло того поворотного момента в истории развития искусственного интеллекта (ИИ), когда создаваемые технологии перешли из области теории в «реальный мир», т.е. стали доступны практически всем отраслям экономики. Этот сдвиг является всеохватывающим, но, в то же время, ставит множество новых проблем, поскольку сочетает в себе сложности применения ИИ со сложностями организации человеческой деятельности, коих отмечается немало. Для адекватного использования новых алгоритмов требуется подготовить целый класс специалистов, способных согласовывать все аспекты жизненного цикла управляемых ИИ производственных и общественных процессов с деятельностью по разработке и совершенствованию научных, экономических и общественных моделей, внедрению их в человеческую деятельность, по их мониторингу и контролю.

ИИ зародился еще в 1950-х годах. Отмечаются две основные эпохи исследований в области ИИ: в 1974–1980 и в 1987–1993 годах, но в то время высокий уровень энтузиазма по поводу применения технологий ИИ так не оправдал ожиданий — были всплески ажиотажа с последующим разочарованием, что приводило к сокращению финансирования исследований, потере интереса к дальнейшему развитию ИИ. Конечно, и сейчас существуют серьезные опа-

сения по поводу того, как будет формироваться будущее ИИ (и его социальных последствий). Краткий ответ будет, скорее всего, такой: из-за гораздо большего (и очень большого) объема актуальных данных и появления процессоров исключительно большой вычислительной мощности (а также графических — GPU) наступления следующей «зимы» искусственного интеллекта ожидать не стоит. Сегодня, при обработке зеттабайтов (ZB) данных, достигается очень высокая точность алгоритмов ИИ в различных сферах человеческой деятельности: от смартфонов до постоянно расширяющейся роли Интернета вещей. В 2020 году в мире было создано более 59 ZB данных. При этом данные генерируются с пятилетним среднегодовым темпом роста в 26 % [1], так что сфера приложений ИИ расширяется невероятными темпами.

ИИ в современном понимании — это довольно широкое понятие, которое на самом деле объединяет несколько различных концепций. В английском языке словосочетание “Artificial Intelligence” не имеет антропоморфной окраски, которую оно приобрело в традиционном русском переводе: слово “Intelligence” в используемом контексте скорее означает «умение рассуждать разумно», а вовсе не «интеллект» (для которого есть английский аналог — “Intellect”). Имеются различные определения искусственного интеллекта как области научно-исследовательской и производственной деятельности общества. Мы в рамках своей сферы интересов так определяем эту область:

«ИИ — это научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными». ИИ — это комплексная дисциплина, являющаяся одновременно формой интеллекта, видом технологии и областью научных исследований, и не предполагающая явного погружения в создание сложных программных кодов.

Сегодня ни у кого не вызывает сомнений, что успехи в развитии экономики страны, в особенности, в её цифровизации, в очень большой степени связаны с внедрением технологий ИИ. В связи с этим возникает принципиально новая задача интеграции ИИ во все отрасли экономики, буквально во все

сферы нашей жизни. Это очень непросто. Перед наукой стоит задача, решить которую в рамках какого-нибудь одного направления нереально. Развитие и повсеместное внедрение технологий ИИ — это, цель, достижимая только серьёзными продвижениями, как в информатике и информационных технологиях, в науке о данных, в машинном обучении, так и в физиологии мозга, психологии, этике, философии и праве. Значимые достижения в этой области могут быть сделаны только объединенными коллективами ученых на стыке разных дисциплин. Но и это — только одна сторона проблемы.

Другая же ее сторона — это нехватка готовых многопрофильных специалистов. Чтобы решать задачи по разработке технологий ИИ, например, в медицинской сфере, информатики должны разбираться в физиологии, а физиологи — в информатике. Мы хорошо понимаем это и считаем, что университетам нужно развивать научно-образовательные программы, по которым студенты будут обучаться всему, что им необходимо для дальнейшей успешной карьеры в выбранной области, а также участвовать в разработке реальных проектов (например, в студенческих конструкторских бюро). Эти задачи сейчас решаются в нашем университете в рамках недавно организованной научно-образовательной школы «Мозг, когнитивные системы и искусственный интеллект», объединяющей группы исследователей и преподавателей ряда факультетов и институтов МГУ имени М. В. Ломоносова.

Мы в своей научно-образовательной деятельности принимаем смену парадигмы в построении образовательной траектории будущих специалистов в области ИИ. А именно, мы считаем, что общество вплотную подошло к восприятию методов и средств ИИ в повседневной жизни как к реализации принципов нового научного знания (учения), базирующегося на следующем фундаменте (основы концептуального подхода авторов изложены в работе [2]).

Основа учения — наука о данных, их представлении, хранении, передаче и анализе, как концептуальной основе моделей жизненных процессов.

Смысл учения — осмысление и представление жизненных процессов в научной, цифровой форме, ориентированной на дальнейший анализ и управление процессами, как в природных, так и в созданных людьми системах развитыми средствами ИИ.

Миссия учения — успешное решение различных задач человеческой деятельности методами и средствами искусственного интеллекта, базирующимися на достижениях науки о данных, как фундаментальной основе.

Миссионеры (научные и технические евангелисты ИИ) — ученые по данным, специалисты в области ИИ.

Средства передачи знаний — высокоразвитые среды обработки данных, такие, как Python с множеством библиотек модулей, а также другие, удобные в использовании средства обработки, представления и анализа данных.

Сфера распространение учения — тотальное обучение представителей всех родов человеческой деятельности основам ИИ, повсеместное внедрение типовых методов и инструментариев. Подготовка ученых и инженеров, осуществляющих развитие и применение средств ИИ.

Постараемся обосновать необходимость нового концептуально восприятия в построении образовательного процесса будущих специалистов. В настоящее время принята следующая классификация специалистов, работающих в области ИИ [3]:

- ученый по данным (Data Scientist);
- инженер по данным (Data Engineer);
- аналитик данных (Data Analyst);
- инженер-программист (Software Engineer).

К этим четырем основным ролям нужно добавить роли, необходимые для переноса задач ИИ в производственную среду:

- специалист по эксплуатации и развитию (DevOps);
- менеджер по управлению рисками / аудитор (Model Risk Manager/Auditor);
- архитектор данных (Machine Learning Architect).

Даже с учетом определенной размытости границ в деятельности этих специалистов можно все-таки утверждать, что, как образовательные стандарты, так и их практическая подготовка должны быть различными. Рассмотрим некоторые аспекты будущей деятельности таких специалистов.

Ученые по данным (Data Scientists).

Потребности специалистов по данным являются наиболее важными — их необходимо учитывать при построении стратегии внедрения ИИ. Специалисты по обработке данных в большинстве организаций сегодня часто имеют дело с разрозненными данными, процессами и инструментами, что затрудняет эффективное масштабирование своих усилий. Хотя большинство экспертов считает роль специалистов по обработке данных в жизненном цикле моделей машинного обучения всего лишь частью построения модели, эта роль должна быть намного шире. С самого начала специалисты по обработке данных должны быть связаны с экспертами в предметной области, чтобы понимать и помогать формулировать бизнес-проблемы таким образом, чтобы они вместе могли создать жизнеспособное решение для ИИ.

Инженеры по данным (Data Engineers).

Конвейеры данных лежат в основе жизненного цикла моделей производственных процессов, а инженеры по обработке данных, в свою очередь, формируют основу конвейеров данных. Поскольку конвейеры данных могут быть абстрактными и сложными, инженеры по обработке данных могут значительно повысить эффективность средств ИИ. В крупных организациях управление потоком данных, помимо управления самими производственными процессами, является постоянной работой.

Аналитики данных (Data Analysts).

Роли ученого по данным и аналитика данных имеют много общего, но для первого обычно требуются более глубокие знания и опыт в области математики (линейная алгебра и многомерное исчисление), а также более продвинутые технические навыки.

Аналитики данных занимаются извлечением соответствующих данных из различных источников и подготовкой их для дальнейшего анализа. На основе анализа аналитик данных должен делать выводы, заполнять отчеты и поддерживать их визуальными эффектами. Наряду с отчетами они должны объяснять, что означают различия в цифрах, если смотреть на них от месяца к месяцу или для разных аудиторий. Основной объем работы аналитиков данных направлен на анализ и описание прошлых или предыдущих стратегий на основе прошлых или текущих данных.

Инженеры-программисты (Software Engineers).

Было бы легко исключить классических инженеров-программистов из рассмотрения области развития методов ИИ, но с более широкой организационной точки зрения важно учитывать их потребности для построения единой корпоративной стратегии внедрения ИИ.

Инженеры-программисты обычно не создают модели управляемых ИИ процессов, но, с другой стороны, большинство организаций создают не только такие модели, но и классическое программное обеспечение и приложения. Важно, чтобы инженеры-программисты и специалисты по обработке данных работали вместе, чтобы обеспечить функционирование более крупной системы. В конце концов, модели ИИ — это не просто отдельные эксперименты — код приложения, обучение, тестирование и развертывание должны соответствовать конвейерам непрерывной интеграции / непрерывной доставки (CI/CD), которые использует остальное программное обеспечение.

Специалисты по эксплуатации и развитию (DevOps).

Развитие средств ИИ базируется на основе принципов DevOps, но это не означает, что они могут работать параллельно как полностью отдельные и разрозненные системы. Команды DevOps выполняют две основные роли в жизненном цикле модели ИИ.

Во-первых, это люди, выполняющие и создающие операционные системы, а также тесты для обеспечения безопасности, производительности и доступности моделей ИИ. Во-вторых, они отвечают за управление конвейером

CI/CD. Обе эти роли требуют тесного сотрудничества с аналитиками данных, инженерами данных и архитекторами данных.

Менеджеры по управлению рисками / аудиторы (Model Risk Managers /Auditors).

В определенных отраслях (особенно в секторе финансовых услуг) функция управления рисками модели (MRM) имеет решающее значение для соблюдения нормативных требований. Но не только строго регулируемые отрасли должны быть затронуты или должны выполнять аналогичную функцию, MRM может защитить компании в любой отрасли от потерь, вызванных плохо работающими моделями ИИ.

Более того, аудит играет важную роль во многих отраслях. Когда дело доходит до жизненного цикла модели ИИ, менеджеры по управлению рисками играют решающую роль в анализе не только результатов использования модели, но и исходных целей и бизнес-вопросов, которые модели ИИ стремятся решить, чтобы минимизировать общий риск для компании. Их следует привлекать вместе с профильными экспертами в самом начале жизненного цикла, чтобы убедиться, что автоматизированный подход на основе ИИ сам по себе не является рискованным.

Архитекторы данных (Machine Learning Architects).

Традиционные архитекторы данных несут ответственность за понимание общей архитектуры предприятия и обеспечение ее соответствия требованиям к данным во всем бизнесе. Обычно они играют роль в определении того, как данные будут храниться и использоваться. Сегодня требования к архитекторам намного выше, они часто должны быть осведомлены не только о тонкостях хранения и потребления данных, но и о том, как модели машинного обучения работают в тандеме. Это усложняет их роль и увеличивает их ответственность в жизненном цикле средств ИИ.

Как мы видим, сфера ИИ сама по себе является многоуровневой, в каждом слое решаются специфические задачи. Поэтому и подготовка специали-

стов по ИИ должна быть выделена в самостоятельную научно-образовательную область, в фундаменте которой находится, прежде всего, наука о данных наряду с примыкающими областями знаний. Реализация этой идеи в том или ином виде осуществляется в ряде ведущих университетов мира (см., например, [4]).

Проблемы обучения студентов «Науке о данных» и подготовки «учёного по данным» имеют свою специфику для классических университетов, готовящих специалистов в рамках программ академического высшего образования. Рассмотрению вариантов ее решения посвящены многие работы российских и зарубежных ученых. Интересно также отметить, что, помимо устоявшихся, сформировавшихся областей, сейчас происходит разработка учебных планов непосредственно для «Науки о данных» [5].

Рассмотрев (пусть даже неполно) спектр направлений деятельности специалистов в области ИИ, мы приходим к выводу, что, несмотря на условность границ их сфер активности, необходимо в процессе обучения учитывать дальнейшую специализацию выпускников. Для различных специалистов необходимо разрабатывать различные образовательные траектории, не пытаясь готовить «универсалов» в области ИИ. Безусловно, в дальнейшем специалисты могут менять профили своей деятельности, но в процессе обучения нужно готовить специалистов различных профилей по самостоятельным образовательным программам. А предметное насыщение этих программ нужно осуществлять с учетом новой учебной парадигмы.

Список литературы

1. *IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data.* URL: <https://www.idc.com/getdoc.jsp?containerId=prUS46286020>.

2. *Главацкий, С. Т.* Наука о данных для математиков – опыт преподавания / С. Т. Главацкий, И. Г. Бурькин. Текст: непосредственный // Современ-

ные информационные технологии в образовании: материалы XXXI международной конференции (Троицк Москва, 2–3 июля 2020 г.). Троицк, 2020. С. 147–149.

3. *Beyond Interactive: Notebook Innovation at Netflix*. URL: <https://netflixtechblog.com/notebook-innovation-591ee3221233>.

4. *Curriculum Guidelines for Undergraduate Programs in Data Science* / De Veaux Richard D., Agarwal Mahesh, Averett Maia, Baumer Benjamin S., Bray Andrew, Bressoud Thomas C., Bryant Lance, Cheng Lei Z., Francis Amanda, Gould Robert, Kim Albert Y., Kretchmar Matt, Lu Qin, Moskol Ann, Nolan Deborah, Pelayo Roberto, Raleigh Sean, Sethi Ricky J., Sondjaja Mutiara, Tiruvilumala Neelesh, Uhlig Paul X., Washington Talitha M, Wesley Curtis L., White David, and Ye Ping // *Annual Review of Statistics and Its Application*. 2017. Vol. 4. P. 15–30. URL: <https://www.annualreviews.org/doi/pdf/10.1146/annurev-statistics-060116-053930>.

5. *Computing Curriculum Committee* / IEEE. URL: <https://www.computer.org/web/peb/curricula>.