

Д.В. Ушаков, академик РАН,  
директор Института психологии РАН

## АЛГЕБРА И ГАРМОНИЯ: ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОЦИОГУМАНИТАРНЫХ НАУКАХ

В наши дни различные научные области специализированы и обособлены, тем не менее наиболее крупные фундаментальные идеи и технологические разработки пересекают дисциплинарные границы и расходятся широкой волной, обогащая значительную часть мировой науки. Социогуманитарные дисциплины с начала нынешнего столетия ощутили влияние по крайней мере двух таких волн. Проект «Геном человека», условным годом завершения которого считается 2003-й, породил, например, палеогенетику, возможность изучать ДНК древних, частично сохранившихся образцов. Благодаря этому далеко вперёд продвинулись наши представления о происхождении человека современного типа и целом ряде событий древней истории, таких как расселение праиндоевропейских племён. В психологии благодаря технологиям, возникшим в результате того же проекта, мы многое узнали о генетической и эпигенетической базе когнитивных процессов и личности, научились точнее определять причины когнитивного дефицита и т.д. Этот ряд можно продолжить и для других социогуманитарных наук.

Вторая волна оказалась ещё более плодотворной, чем первая, и связана с искусственным интеллектом (ИИ) – «революцией глубокого обучения» (начало которой условно относят к 2012 г., когда нейросеть смогла победить на конкурсе машинного зрения), и с появлением больших языковых моделей в конце 2022 г. Можно выделить три большие сферы, где социогуманитарные науки оказались под сильным влиянием событий в области искусственного интеллекта.

Во-первых, искусственный интеллект даёт мощные средства выявления различных закономерностей, в том числе закономерностей поведения человека и функционирования общества. Эти средства порой превосходят человеческие возможности и увеличивают скорости роста научного знания и технологий, чем стали успешно пользоваться социогуманитарные науки.

Во-вторых, продолжающееся развитие технологий искусственного интеллекта всё сильнее влияет на человека и общество, что ставит перед социогуманитариями сложную задачу оценки этого влияния, где-то – его предвосхищения и даже проектирования.

Наконец, создание искусственного интеллекта всегда в большей или меньшей степени стимулировалось знаниями о том, как устроены человеческий мозг и психика. И обратно – возникновение искусственных когнитивных систем давало ценный материал для исследования систем естественных. По мере прогресса технологий искусственного интеллекта и наших знаний о мозге и психике это взаимодействие возрастает.

Эти три сферы в той или иной степени отражены в предлагаемом вниманию читателя специальном выпуске журнала.

### Общество сосуществования естественного интеллекта с искусственным

Где-то со второй половины XX в. человечество вступило в период, когда стали просматриваться перспективы усиления интеллектуальных возможностей за счёт технологий двух типов. Во-первых, это по-

вышение биологического разума, то есть создание технологическим путём человека сверхразумного. Во-вторых, конструирование искусственного сверхразума.

Какое-то время назад первый путь казался более реалистичным. Геном человека расшифрован, и уже работают методы редактирования генома, такие как CRISP. Тем не менее сегодня внесение изменений в человеческий геном находится под строгим этическим запретом, а перспективы биологического изменения человека и увеличения его интеллекта весьма туманны, хотя и полностью отрицать такую возможность нельзя.

А вот в другом направлении – в сфере искусственного интеллекта – в последнее десятилетие произошёл колоссальный рост, что заставило всерьёз задуматься о перспективах появления так называемого сильного искусственного интеллекта, то есть такого, который во многих или даже в большинстве проявлений сильнее человеческого разума. Туманные контуры сильного искусственного интеллекта на горизонте обеспокоили и вызвали к жизни прогнозы и призывы со стороны таких влиятельных людей, как Б. Гейтс, С. Хокинг, И. Маск, лидеры крупнейших ИТ-корпораций. Они сравнили перспективы человечества, существующего с сильным искусственным интеллектом, с судьбой муравейника, который затапливают при строительстве плотины не из-за желания зла, а просто в связи с необходимостью решения крупных задач.

Всё же восстание машин, к счастью, выглядит сомнительным в эволюционной перспективе. Для восстания машинам требуется обладать не только выдающимися когнитивными возможностями, но и мотивационной системой, завязанной в том числе на гомеостатические потребности тела, а также выходом на материальные и идеальные инструменты человеческой деятельности. Воспроизведение всех этих элементов фактически означает сотворение искусственной жизни, причем в её разумной форме.

Из сказанного возникает видение наиболее желательной перспективы: это существование людей, придающих смыслы происходящему, с сильным искусственным интеллектом, превосходящим их в вычислительных возможностях. Такое видение перспектив заставляет задаться вопросом, как человеку оказаться субъектом, то есть активным по отношению к этому искусственному интеллекту? Как избежать использования новых мощных технологий для манипуляций и закабаления одних людей другими?

Визионерские оценки дальних перспектив эволюции разума на Земле в последнее время дополняются вполне конкретным обсуждением и решением проблем, связанных с развитием технологий искусственного интеллекта уже сейчас или в ближайшие годы. Различные стороны этих проблем рассмотрены в статьях, публикуемых ниже.

Правовое измерение формирования новой техносферы анализируется в статье академика РАН Т.Я. Хабриевой. Рассматривая национальные, наднациональные и международные уровни регулирования ИИ, автор показывает, как складываются различные регулятивные подходы – человекоцентричный, ориентированный на защиту прав и свобод, и подход технологического лидерства, делающий ставку на ослабление ограничений ради ускорения инноваций. Особый интерес представляют критический анализ Глобального цифрового договора и обсуждение рисков формирования новой системы глобального управления, в которой принципы равенства и справедливости оказываются под давлением экономических и geopolитических интересов.

В статье И.В. Данилина анализируются экономические и политические вызовы распространения ИИ, в частности вопросы структуры цифровых рынков, монополизации ключевых сегментов крупными корпорациями, а также сложности регулирования в условиях секьюритизации технологической повестки и geopolитиче-

ской конкуренции. Разработка и владение искусственным интеллектом всё больше становятся факторами, определяющими власть и могущество. Лидирующую роль в этих разработках, как известно, играют технологические корпорации всего нескольких стран мира. Запрос на вычислительные мощности, соответствующую им энергетику и базы данных для обучения, потенциальная возможность роста безработицы – всё это проблемы, встающие сегодня перед человечеством.

Социологический и философский курс представлен в работе В.К. Левашова и О.В. Гребняк, где обсуждается судьба гуманистических принципов и ценностей в контексте перехода к цифровой цивилизации. Опираясь на концепцию ноосферы и данные российского социологического мониторинга, авторы показывают амбивалентность цифровых технологий. С одной стороны, они открывают возможности решения глобальных проблем и оптимизации управления, с другой – усиливают неравенство и противоречия, фрагментируют информационное пространство и подрывают доверие к информации. В центре обсуждения – необходимость гуманитарной экспертизы и этических ограничителей, без которых цифровая трансформация обрачиваются дегуманизацией.

В контексте развития глобальных технологий искусственного интеллекта возникает проблема сохранения культурной специфики больших и малых народов. В статье Д.И. Коломацкого и Е.В. Коровиной на материале проектов из разных регионов мира показано, что внимание к миноритарным языкам может означать как поддержку, так и новую форму эксплуатации, когда сообщества рассматриваются прежде всего как источник данных. Авторы настаивают на том, что участие носителей, учёт их прав и интересов, прозрачность целей исследований являются необходимым условием этически оправданного применения ИИ.

Статья Ю.В. Балакиной, М.В. Григорьевой, А.Г. Смирнова и Е.Н. Соколовой по-

священа возможностям русскоязычных языковых моделей в работе с ценностно ориентированным контентом. Авторы исходят из того, что для русского языка характерны не только особая морфология и сложная синтаксическая организация, но и специфический ментальный контекст, уходящий корнями в философию и педагогику XIX–XX вв. В условиях сетевой коммуникации, где разговорность, ирония и спонтанность меняют структуру обыденного языка, становится особенно значимой задача нравственно-семиотического моделирования – разработки моделей, которые учитывают не только семантику, но и ценностный код русской культуры. Тем самым ИИ рассматривается как участник ценностной коммуникации, а не только как статистический предсказатель текста.

### **Искусственный интеллект как средство познания и помощи человеку и обществу**

Технологии искусственного интеллекта к сегодняшнему дню научились распознавать скрытые закономерности в данных. Сказанное относится и к данным, касающимся человека и общества. Так, в последние несколько лет разработаны средства автоматического распознавания свойств человека, таких как Большая пятёрка личностных свойств или когнитивные способности. Искусственный интеллект способен достаточно успешно идентифицировать и психологические состояния людей. Эта тема имеет большое практическое значение в разных областях – для требующих напряжения профессий, для персонала корпораций, в школе, для оценки психологического состояния больших социальных групп, регионов и т.д. Распознавать индивидуальные особенности и состояния можно на основании разных типов информации: видеointerview в стандартизованных условиях, активности в социальных сетях, психофизиологических данных, таких как КГР, ЭЭГ, оценки сердечного ритма и т.д.

Способность к распознанию закономерностей позволяет не только идентифици-

ровать состояния и свойства, но и прогнозировать поведение. В связи с этим активно развивается тема цифровых двойников человека, которые на основе заложенной в них информации о человеке предсказывают его поведение в тех или иных ситуациях. Для этих целей оказалось эффективным использование больших языковых моделей, которые фактически содержат в себе культурный опыт человечества, зафиксированный в текстах. Общее моделирование человеческих аттитюдов и поведения позволяет предсказать, исходя из специфических сведений о человеке, его поведение в широком круге ситуаций – от приобретения товаров и политической позиции до семейной жизни.

Статья А.Р. Бахтизина и А.В. Брагина, публикуемая ниже, хорошо показывает возможности использования ИИ как инструмента макроэкономического прогнозирования. Хотя большие языковые модели ещё весьма новы, краткосрочные прогнозы (до одного года) на их основе уже сейчас оказываются эффективными. В то же время по мере увеличения глубины прогнозирования ИИ начинает уступать другим инструментам. Авторы сравнивают преимущества и ограничения эконометрических, структурных и агент-ориентированных моделей и современных трансформерных архитектур. Показано, что ИИ способен интегрировать разнородные данные и улучшать точность прогнозов, но сталкивается с проблемами интерпретируемости и устойчивости. Предлагаемая перспектива гибридных подходов – сочетание вычислительной мощности ИИ с теоретической строгостью экономических моделей – перекликается с идеей всего выпуска о необходимости диалога между умными технологиями и дисциплинарными традициями знания.

В статье Л.И. Бородкина прослеживается эволюция применения ИИ в исторических исследованиях – от первых экспериментов с экспертными системами до современных генеративных моделей. Особое внимание уделяется распознаванию рукописных

и старопечатных текстов, реконструкции утраченного культурного наследия, а также методологическим и этическим вопросам использования генеративных нейросетей и чат-ботов. Автор обращает внимание на повышение доступности источников и возрастающий риск квазизнаний, когда убедительно выглядящие ответы системы подменяют собой реальную мысль.

Статья А.В. Сиренова, Е.В. Анисимова, Т.А. Базаровой и М.Е. Проскуряковой показывает, как проблемы решаются в конкретном крупном проекте. Совместная работа историков, палеографов и специалистов по анализу данных над распознаванием автографов Петра Великого привела к созданию модели, которая считывает сложный почерк с высокой точностью и делает петровские письма доступными не только узкому кругу палеографов, но и широкой аудитории. Через эту оптику ИИ предстаёт как инструмент демократизации доступа к письменному наследию и одновременно как повод для переосмысления принципов научной атрибуции и публикации источников.

Технологии искусственного интеллекта помогают не только познавать человека, но и оказывать ему разностороннюю поддержку. Одна из сфер, где такая поддержка востребована, – это образование, в частности – разработка индивидуальных образовательных траекторий. Машинное обучение на массивах данных по ученикам, осваивающим тот или иной учебный курс, позволяет предсказывать, что способен освоить или решить тот или иной ученик при определённом профиле предшествующих достижений. Однако работа с цифровыми двойниками может дать ещё больше: специально сконструированные для ученика задания на основе его когнитивного развития в соответствующей области. Пример реализации такого подхода можно уже сегодня увидеть в области тестов интеллекта, генерируемых компьютером.

Статья С.Ю. Степанова обращается к проблеме искусственного интеллекта как личного помощника и друга челове-

ка. Прослеживая путь от классических представлений о рефлексии до формирования психологии рефлексии и современных рефлексивных практик, автор вводит понятие би-рефлексии – взаимодействия человеческого сознания и машинного интеллекта в гибридных конфигурациях цифровых кентавров. Концепты Цифрового ангела и Цифрового альтер-эго описывают возможные форматы персональных ИИ-агентов, которые поддерживают саморегуляцию, расширяют пространство саморазвития и в то же время ставят острые вопросы о границах идентичности и ответственности человека в цифровой среде.

### **Параллели между естественным и искусственным интеллектом**

Между созданием искусственных нейронных сетей и исследованиями человеческого мозга существуют глубокие исторические взаимосвязи, продолжающиеся сегодня нарастающим междисциплинарным взаимодействием специалистов. Корни идеи нейронных сетей уходят вглубь времён, предшествующих наукам об искусственном интеллекте, в философско-психологический ассоциализм. В конце XIX – первой половине XX в. к работе подключились и физиологи. Благодаря трудам И.П. Павлова, В.М. Бехтерева, Э. Торндайка, Б.Ф. Скиннера и др. идеи ассоциации распространились на поведенческие акты и обучение. В 1949 г. Д. Хебб наложил эти идеи на представления о нейронах, введя гипотетический синапс, формирующий связи при совпадении разрядов, – первый алгоритм обучения и шаг к распределённому хранению знания в группах нейронов. Нейрофизиолог и психолог Ф. Розенблatt сконструировал Персептрон (1957 г.) – первую нейросетевую модель со способностью к распознаванию образов.

Именно нейросетевые модели, подготовленные длительным развитием идей в философии, психологии и нейрофизиологии, лежат в основе современных технологий ИИ. При этом искусственный

интеллект в его нынешнем виде работает совсем не так, как человеческий мозг. Нет и полного междисциплинарного единства в понимании интеллекта: там, где психологии и физиологи следуют фактам, ИИ-специалисты стремятся к эффективности.

Например, обратное распространение ошибки в искусственных нейросетях (Галушкин, Вербос и др.) не биологично, поскольку в мозге не обнаружены требуемые таким механизмом нейронные пути. А синапс Хебба, который находит подтверждение с помощью оптогенетики, не применяется в искусственных нейросетях. Другой пример: Нобелевские лауреаты 1981 г. Д. Хьюбел и Т. Визель открыли нейроны-детекторы линий и вдохновили на создание свёрточных сетей. Однако, как показывают новые исследования, лишь 10% клеток зрительной коры соответствуют этим принципам.

Вся эта сложная ситуация делает непростым, но продуктивным междисциплинарное взаимодействие специалистов в области естественного и искусственного интеллекта. Эта тема заслуживает отдельного рассмотрения. В настоящем специальном выпуске размещена статья Т.О. Шавриной и А.А. Корнилова о мультиязычности в языковом моделировании, которая показывает, как специалисты в области искусственного интеллекта могут использовать лингвистические знания. Авторы предлагают систематический обзор поддержки малоресурсных языков в современных языковых моделях и демонстрируют подход на основе дополненной генерации, который позволяет использовать дескриптивные грамматики и типологические базы данных для решения задач перевода и анализа языков, фактически не представленных в цифровом пространстве.

В целом материалы, публикуемые в номере, позволяют заключить, что проблематика искусственного интеллекта находится в самом центре современного социогуманитарного знания.