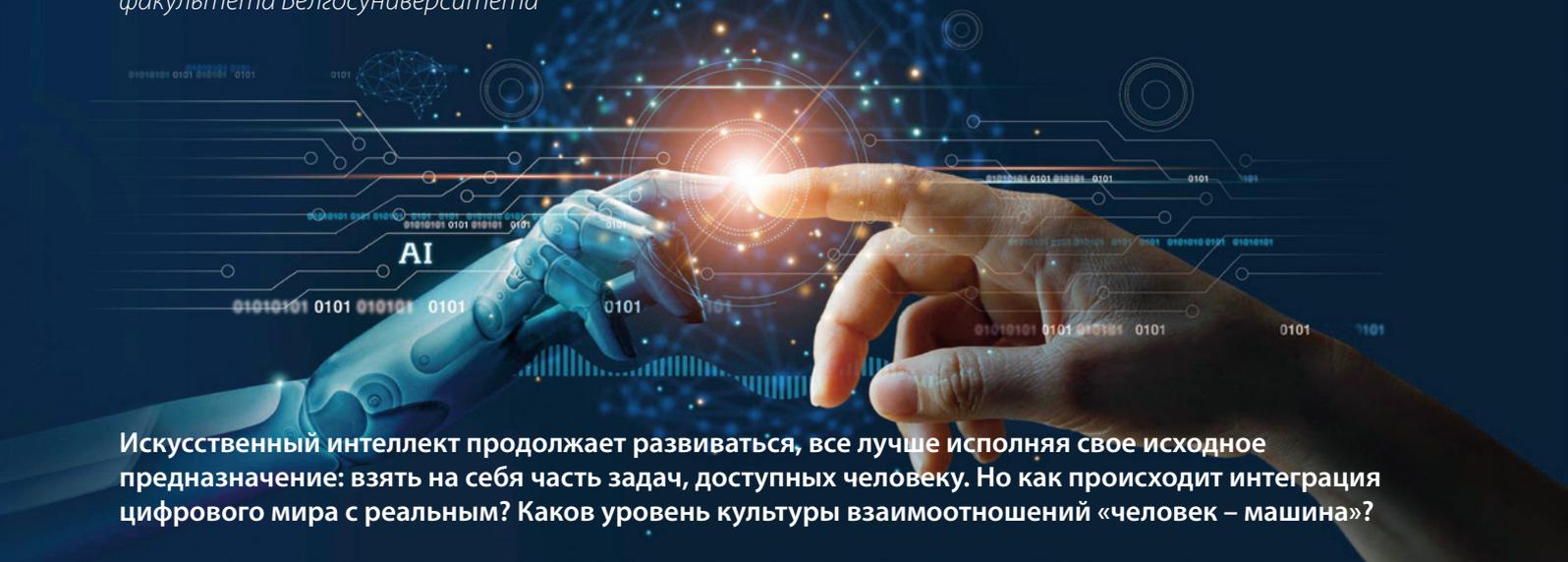


ИИ и КУЛЬТУРА: запрещать или регулировать?

Борис ПАНЬШИН,

доктор технических наук, профессор экономического факультета Белгосуниверситета



Искусственный интеллект продолжает развиваться, все лучше исполняя свое исходное предназначение: взять на себя часть задач, доступных человеку. Но как происходит интеграция цифрового мира с реальным? Каков уровень культуры взаимоотношений «человек – машина»?

ПРЕДТЕЧИ

В 2023 и 2024 годах научная общественность будет отмечать 100-летие выдающихся советских ученых – Виктора Михайловича Глушкова и Виктора Владимировича Мартынова. Первый был одним из самых гениальных и выдающихся людей XX столетия, «царем советской кибернетики». Первая в мире «Энциклопедии кибернетики» была подготовлена по инициативе Глушкова. Он предусматривал, насколько важную роль в жизни общества будет играть вычислительная техника.

В. В. Мартынов – основоположник белорусской школы формализованной семантики, заслуженный деятель науки БССР, доктор филологических наук, профессор. Сфера его научных интересов простиралась от литературоведения, культурологии, сравнительно-исторического языкознания, этимологии, славистики, теории языка и востоковедения до информатики, искус-

ственных интеллектуальных систем и технологий.

Работы этих ученых в области кибернетики и лингвистики не теряют актуальности и имеют фундаментальное значение для создания искусственного интеллекта (ИИ) будущего.

Разработанный белорусским ученым-лингвистом еще в 1960–1980-х годах прошлого века Универсальный Семантический Код (УСК – способ «исчисления языковых смыслов») до сих пор уникален и занимает особое положение в представлении и преобразовании знаний. Можно предположить, что разработанные В. В. Мартыновым алгебры УСК применяются в современных алгоритмах обучения нейросетей. Они не имеют аналогов в формальном представлении и преобразовании смыслов, и альтернативу им трудно представить.

Примечательно, что в тот же период другим выдающимся ученым,

кибернетиком В. М. Глушковым, были разработаны алгебры теории цифровых автоматов и практики проектирования на их основе ЭВМ и создания языков программирования.

Можно предположить, что объединение этих разработок неизбежно и естественно привело бы к созданию первого в мире семантического компьютера (УСК-машины), что могло позволить уже в то время решать сложные интеллектуальные задачи и создавать фильтры для борьбы с нарастающей угрозой «загрязнения» информационной (цифровой) среды в условиях семантической неопределенности создаваемых и наполняющих сеть данных. Одновременно упростились бы задачи подготовки данных для создания общегосударственной автоматизированной системы (ОГАС). Сказался недостаточный уровень междисциплинарности исследований.

В настоящее время работы по системам ИИ успешно развиваются

в БГУИР, БГУ, БГЭУ, других вузах страны и в Межведомственном центре по искусственному интеллекту НАН Беларуси.

ИНТЕРФЕЙСЫ НОВОГО ВРЕМЕНИ

Четкого и однозначного определения понятия «искусственный интеллект» нет. Тем не менее ИИ стал не только модным названием, он уже проник во многие сферы деятельности отдельного человека и целых предприятий, использующих современные информационные и цифровые технологии.

В общем виде ИИ – это широкая область науки и технологий, предназначенных для решения отдельных интеллектуальных задач. В виду глобальности интернета и роста масштабов применения цифровых технологий ИИ можно определить как главную технологию XX века.

В практическом плане ИИ – это способность автоматизированных систем понимать и брать на себя отдельные функции интеллекта человека (например, выбирать и принимать решения на основе ранее полученного опыта и рационального анализа внешних воздействий). Основными направлениями использования ИИ являются: разработка естественно-языковых интерфейсов и машинный перевод, исследование методов и разработка систем для общения человека с компьютером на естественном языке, генерация и распознавание речи, обработка визуальной информации, сервисы по оцифровке документов (актуальность договоров, наличие необходимых подписей, печатей, дат и реквизитов) и многое другое для помощи человеку в интеллектуальной работе.

С рутинной и объемной работой нейросети справляются гораздо лучше людей, которых долгий монотонный труд приводит к выгоранию и изнуряет. По сути, технологии ИИ становятся обычными инструментами в сборке, систематизации и анализе информации для принятия управленческих решений и проведения компьютерных экспериментов вместо натуральных испытаний.

МОЖНО СКАЗАТЬ, ЧТО ИИ – ЭТО ОБЛАСТЬ, КОТОРАЯ СОЧЕТАЕТ В СЕБЕ НАДЕЖНЫЕ НАБОРЫ ДАННЫХ ДЛЯ РЕШЕНИЯ ПРОБЛЕМ, АЛГОРИТМЫ И МОЩНЫЕ КОМПЬЮТЕРЫ, ТО ЕСТЬ: ИИ = БОЛЬШИЕ ДАННЫЕ + АЛГОРИТМЫ + МОЩНЫЕ СУПЕРКОМПЬЮТЕРЫ

ОСТРОТА ПРОБЛЕМЫ – В УРОВНЕ КУЛЬТУРЫ РАЗРАБОТЧИКОВ

Актуальность разработки и применения технологий ИИ обусловлена ростом сложности объективно необходимых задач управления экономикой и постоянно возрастающими темпами и масштабами цифровизации.

Предполагается, что ИИ поднимет мировой ВВП на 7 % и автоматизирует до четверти рабочих мест в развитых странах. В России, по расчетам Сбера, мультипликативный эффект от применения ИИ-решений к 2025 году может добавить к ВВП страны минимум 1 %. Для Беларуси эта оценка может быть значительно выше вследствие большей организованности и связности экономики (все отрасли работают как единый народнохозяйственный организм, а не как отдельные бизнесы).

Как правило, технологии ИИ наиболее активно используются в постцифровых секторах экономики (финансы, электронная, гейминг), где основной массив данных изначально оцифровывается, так как они собираются в электронном виде. При этом наибольший экономический эффект приносит вовсе не «бытовые» ИИ-решения (например, разговорные боты в колл-центрах), а системы, направленные на поддержку профессиональной деятельности (например, банковский скоринг или маркетинг – предсказание будущих покупок клиентов).

В настоящее время ИИ входит в новую фазу развития и все более становится одним из главных катализаторов изменений в экономике и образовании. В то же время решения, как использовать эту тех-

нологию, уравнивать риски и возможности, принимают в первую очередь крупные корпорации, оттесняя на второй план исследования по оценке рисков разработки и последствий внедрения приложений с технологиями ИИ для органов государственного управления и предприятий реального сектора экономики. Как следствие, развитие ИИ идет очень противоречиво и зигзагообразно. Крупные компании включились в гонку за быстрейшую монетизацию технологий ИИ. Если в 2022 году интернет-гигантами и крупными корпорациями было создано 32 значимые модели машинного обучения, то научными центрами только три.

Одновременно наблюдается тенденция сокращения отделов, занимающихся вопросами этики и безопасности, даже в таких корпорациях, как Microsoft или Google. В основном это происходит из-за того, что развитие технологий ИИ стало требовать все больше ресурсов: кадровых, информационных (базы данных) и вычислительных мощностей, необходимых для создания таких приложений. Как следствие, возрастает интерес к регулированию сферы ИИ со стороны государственного управления: анализ ситуации в 127 странах показал, что количество принятых в разных странах законов, содержащих словосочетание «искусственный интеллект» выросло с двух в 2016 году до 37 в 2022-м. В экспертной среде все более активно обсуждается необходимость скорейшего внедрения и актуализации этического кодекса искусственного интеллекта на национальном и международном уровнях.

В целом вопросы, что можно доверить искусственному интеллекту,

а что категорически нельзя, связаны с общим уровнем культуры разработчиков и пользователей ИИ технологий и цифровой культуры, в частности.

СИНТАКСИС И СЕМАНТИКА ИИ – ЗЕРКАЛО КУЛЬТУРЫ ЦИФРОВОЙ СРЕДЫ

Интеллект, как естественный, так и искусственный, предполагает способность понимать и соединять некоторые смыслы (понятия, гипотезы, концепции, алгоритмы, научные теории и т. д.), представленные в виде знаков, объединенных по соответствующим грамматическим правилам. То есть в общем виде, речь идет о семантике (смысле) и формах знакового представления смыслов (синтаксисе).

Отсюда следует, что для понимания сути технологий ИИ необходимо различать знаковое мышление, которое оперирует не самими смыслами, а знаками этих смыслов (буквами, словами, предложениями), и семантическое мышление – оперирование смыслами, представленными в виде знаков.

Синтаксис (грамматика для моделей ИИ) – это законы организации знаков, которые можно формализовать, и затем исчислять по определенным алгоритмам, которые представлены, как правило, в алгебраической (символы элементов и операций с ними) или графической форме (графы, представленные узлами и дугами) и затем реализуются с помощью компьютерных программ – семантических машин (СМ). Ярким примером таких СМ являются электронные переводчики текстов, чат-боты, генераторы речи, звуков и изображений.

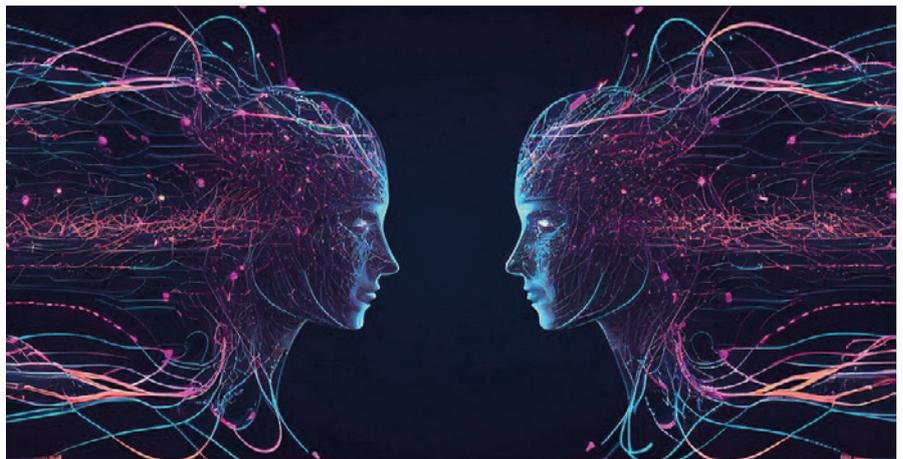
Семантика – это понимание того, как знаковые формы соединяются с содержанием. Здесь уже в плане формализации все кардинально усложняется, и пока имеются только достаточно простые системы, функционирующие по алгоритмам обработки текстов, звуков изображений, характеризующихся достаточно видимой логикой их представления. Например, своды юридических правил или известные системы ТРИЗ.

Но это не мышление, а имитация мышления по заранее определенным правилам. Новой информации и новых знаний здесь не возникает. Поэтому современные технологии ИИ правильнее называть имитирующими человеческое мышление технологиями. В сфере ИИ пока освоены семантические машины, работа которых (например, ChatGPT) создает для человека иллюзию мышления, так как результаты, выдаваемые компьютером, человек соединяет с имеющейся у него картиной мира. Отсюда возникает иллюзия ИИ (якобы машина мыслит), но на самом деле никакого мыслительного процесса не происходит. Здесь применены алгебры и вычислительные процедуры работы с синтаксисом по определенным жестким правилам, исполняемым мощными компьютерами. Это имитация мышления, за которой всегда стоят люди.

Появление и широкое распространение чат-ботов типа GPT – это революция в сфере знаковых систем, обусловленная: новыми, более развитыми алгоритмами работы со знаковыми системами, которые ранее не удавалось формализовать; мощными суперкомпьютерами; громадными массивами оцифрованных данных (Big Date). Главное – переформировать интеллектуальную задачу в виде текста, который затем можно продолжить. Такой ИИ эффективен при решении отдельных узкоспециализированных задач, для которых имеются оцифрованные массивы откорректированных данных. Поэтому важна культура подготовки первичных данных. Любая задача может быть описана, и любая задача может быть решена путем продолжения текста. Случайностей в природе нет, все предопределено необходимостью и алгоритмами, о которых мы пока еще не знаем.

НЕЙРОСЕТЬ – СЛОЖНАЯ И БОЛЬШАЯ МНОГОЭТАЖНАЯ ФОРМУЛА, ГДЕ ПОНЯТНА КАЖДАЯ ОПЕРАЦИЯ, НО НЕ ПОНЯТНО ИХ КОЛЛЕКТИВНОЕ ДЕЙСТВИЕ.

ОБУЧЕНИЕ ЗАВИСИТ ОТ МНОГИХ ДАННЫХ. ИЗ-ЗА АЛГОРИТМИЧЕСКИХ ДЕФЕКТОВ ВОЗМОЖНЫ ОШИБКИ. МЫ ВСЕ – ЧЛЕНЫ ОБЩЕСТВА – СОЗДАЕМ ЯЗЫК И НАБОРЫ ДАННЫХ, ПОЭТОМУ ВАЖНЫ КУЛЬТУРА В ЕЕ ТРАДИЦИОННОМ ПОНИМАНИИ И КУЛЬТУРА ЦИФРОВАЯ, УЧИТЫВАЮЩАЯ СПЕЦИФИКУ ЦИФРОВОЙ СРЕДЫ.



Для обучения нейросетей создаются наборы тестов, чтобы определить ее возможности решать конкретные задачи. Пока лучшие результаты получены для среднего уровня знаний.

ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ И ПЕРСПЕКТИВЫ ИИ

На очереди – следующая революция в сфере технологий ИИ, то есть создание моделей глубокого машинного обучения. Общее направление – универсализация фундаментальных моделей машинного обучения. Если такую модель можно будет обучить на правильных текстах, то нейросеть сможет решать практически любую задачу. Это займет еще немало времени, но добиться даже средней сопоставимости ИИ с естественным в обозримой перспективе нереально и вряд ли будет возможно, так как особенность человеческого интеллекта заключается в его универсализации и колоссальном быстродействии (86 млрд нейронов, квадриллион синапсов). Суперкомпьютеров, сопоставимых с мозгом, пока не создано.

НЕ НАДО ИЛЛЮЗИЙ НА ФОНЕ УГРОЗ

Технологии в сфере искусственного интеллекта появляются быстрее, чем мы успеваем их осмысливать. Возникают иллюзии, что если у нас будет достаточно данных и мощных компьютеров, то мы сможем поручить ИИ любую задачу. Однако пока основные эффекты связаны с дипфейками и игровым контентом. Не секрет, что различные технологии больших данных и машинного обучения используются для манипулирования массовым сознанием, а дипфейк – как раз один из мощных инструментов для подобных действий.

По данным компаний, специализирующихся на цифровой безопасности, до 80 % киберпреступлений совершается с помощью социальной инженерии (то есть дипфейков), а не уязвимостей в защите компьютеров.

В 2022 году ИИ перешел в «стадию развертывания»: технологии быстро прогрессируют, появляются все новые модели и коммерческие продукты. В процесс включены как крупные компании, так и стартапы – технология ИИ больше не контролируется небольшой группой лиц. В 2022 году и в начале 2023 года регулярно выпускались новые масштабные модели ИИ, способные выполнять все более широкий круг задач – от анализа текста до создания изображения и качественного распознавания речи (примером является ChatGPT). Глобальные угрозы несет простота создания дипфейков, которые, подобно пеплу Везувия, скрывшему Помпею, накроют Сеть дипфейками и информацией Second Hand. Так как нейросети обучаются на основе уже созданных данных, новая информация не возникает, а без этого нет развития.

В результате, как отмечают многие эксперты, перед мировым сообществом встают вопросы об этических проблемах, связанных с развертыванием ИИ и мерой его регулирования. Странам предстоит решить, надо ли приводить использование ИИ к человеческим этическим нормам, что сильно ограничит технологию и снизит ее эффективность, или в интересах конкуренции и развития дать разработчикам больше свободы. По сути, речь идет о цифровой культуре ИИ.

Сегодня разработка ИИ по-прежнему вызывает много ажиотажа, однако рынок ИИ вступает в период охлаждения, свидетельством чему является снижение инвестиций в разработку ИИ-технологий.

Как следует из недавнего отчета Стэнфордского университета, в 2022-м рост частных вложений в технологии ИИ в мире приостановился после десяти лет роста. В 2022 году частные инвестиции в ИИ в мире составили более 91 млрд долл. США, на 27 % меньше, чем в 2021 году, когда показатель составлял рекордные 125 млрд долл. США. Сокращение показателя в 2022 году – первое за последнее десятилетие.

Лидером частных вложений в ИИ остаются: США – 47 млрд долл. США, Китай – 13 млрд долл. США, ЕС и Великобритания – 11 млрд долл. США. По числу стартапов в области ИИ США занимают первое место в мире – 542. В Китае – 160, в Великобритании – 99. В 2022 году наибольший объем инвестиций был направлен в такие сферы, как медицина и здравоохранение (6 млрд долл. США), обработка и управление данными и облачные технологии (6 млрд долл. США), финтех (5 млрд долл. США).

Сегодня начинают обсуждаться вопросы, которые выходят за пределы самих технологий ИИ и касаются развития человеческого сообщества в целом. Выявляемые ошибки и проблемы при использовании ИИ указывают на этические проблемы регулирования этой сферы.

К примеру, Microsoft впервые ограничил доступ к своим технологиям ИИ – распознаванию изображений и синтезированию голосовых подделок (дипфейков). Это произошло из-за принятия нового стандарта этики в отношении ИИ, согласно которому интересы личности должны стоять выше технологий. Влияние ИИ на реальную жизнь активно обсуждается российскими и белорусскими учеными.

Очевидно, что с ростом опасений по влиянию ИИ на рынок труда регулирование применения ИИ будет совершенствоваться по различным направлениям, что требует разработки соответствующей методологической основы для оценки рисков ИИ и выбора направлений развития и применения приложений с технологиями ИИ. В части направлений, по нашему мнению, – это экономика и образование. В части методологии – применение принципов синергетики к построению качественных моделей (фазовых портретов) развития экономики под влиянием роста применения ИИ.

ЦИФРОВАЯ КУЛЬТУРА ДЛЯ ИИ

Возникновение цифровой культуры обычно связано с набором практик, основанных на все более интенсивном использовании цифровых технологий.

Ключевой является проблема цифрового доверия как нового явления цифровой эпохи, которое приобретает стремительно возрастающую актуальность наряду с кибернетической и информационной, а также цифровой экономической безопасностью как отдельного человека, предприятия, отрасли, государства и сообщества стран в целом. Прогнозируется, что к 2025 году кибератаки будут причинять ущерб в размере 10,5 трлн долларов США в год. По мнению многих экспертов, это только вершина айсберга, так как большая часть атак (по разным оценкам до 80 %) остается неизвестной. При этом в 90 % случаев информационные системы компаний могут быть взломаны в течение от 1 до 5 дней. С применением технологий ИИ угрозы и риски взлома сайтов и аккаунтов резко возрастают. Потери мировой экономики от дипфейков уже в 2021 году оценивались более чем в 400 млрд долл. США. И, очевидно, что эти потери будут возрастать вместе с ростом предвзятости алгоритмов генерации текстов и поисковых систем.

Следовательно, для цифровой сферы необходимо ускорить формирование соответствующего культурного и нормативно-правового обеспечения. Недавно опубликованные Пекинские принципы искусственного интеллекта и аналогичные принципы, разработанные в ЕС, России и ряде других стран демонстрируют существенное совпадение оценок и концепций реакций на ключевые

проблемы современного состояния с применением систем ИИ.

Для сферы ИИ нужны люди, специально разрабатывающие технологии понимания ИИ как область сценариев, которая программирует человека на использование различных интерфейсов и сервисов. Ключевым звеном становится уровень цифровой культуры программиста и оператора подготовки первичных данных и текстов.

Поэтому ключевое значение имеет соответствующая подготовка и переподготовка кадров. Предполагается, что интеллектуальные цифровые технологии вытеснят людей в рутинных задачах, а люди будут успешны в занятиях, требующих хороших социальных навыков (soft skills) и междисциплинарного опыта.

Становится все более очевидным, что внедрение технологий ИИ в отраслевую деятельность (в первую очередь в промышленности) является определяющим фактором конкурентоспособности. Тот, кто не использует ИИ, не может играть на рынке на равных (опаздывает по срокам, проигрывает по себестоимости). По оценкам PwC, Соединенные Штаты и Китай намерены получить как минимум 70 % от 15,7 трлн долл. США, которыми ИИ обогатит мировую экономику к 2030 г., причем на долю Китая придется 7 трлн долл. США.

В цифровой экономике важно не только то, с какой скоростью действуют первичные элементы (экономические агенты, находящиеся

в постоянных взаимодействиях друг с другом), но и особенности тех связей и отношений, в которые вступают эти агенты (обезличенные или личные, подверженные субординации или самостоятельные и т. д.). То есть все то, что невозможно или нецелесообразно детально прописать в должностных инструкциях и что должно реализоваться в режиме самонастройки, достижение которого осуществляется путем повышения уровня организационной и цифровой культуры персонала и частного лица в соответствии с принципом ответственности.

Вследствие самонастройки происходит замещение механизмов экономического и административного принуждения механизмами самоконтроля и саморегуляции персонала и всей производственной системы. В этом и состоит суть влияния организационной и цифровой культуры на снижение трансакционных издержек и повышение экономической эффективности производственной системы в целом.

Цифровая культура искусственного интеллекта усиливает значение традиционной организационной культуры предприятия. Она все более становится очевидной характеристикой современной экономической системы и начинает увереннее включаться в экономический анализ эффектов цифровой трансформации.

Можно сказать, что с ростом применения технологий ИИ в интернете наступает приоритет культуры, а не экономики. Только программного обеспечения для защиты цифрового (информационного) пространства от современных угроз цифровизации недостаточно. Настала пора на государственном и международном уровне принимать социально одобряемые нормы поведения в Сети и применения ИИ-технологий. В конечном итоге именно на таком подходе должны базироваться новые современные правила коммуникации.

