ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПРОШЛОЕ, НАСТОЯЩЕЕ, БУДУЩЕЕ

Часть IV

Юрий ВИЗИЛЬТЕР,
начальник подразделения интеллектуального анализа данных и технического зрения ФГУП «ГосНИИ Авиационных систем», доктор физико-математических наук, профессор РАН

Современной революции искусственного интеллекта уже 10 лет (2011-2021). Что ожидает нас в следующее десятилетие? В первой части статьи (в №6-2020 журнала «Авиапанорама») мы начали разговор о том, что мировая реальность 2020-х годов во многом будет определяться уровнем достижений в области искусственного интеллекта (ИИ). И в этой изменившейся реальности России потребуются технологическая независимость и технологический паритет в сфере интеллектуальных технологий. В заключительной части статьи мы поговорим о том, какие задачи, проблемы и вызовы будут стоять перед нами в ближайшие годы. Начнем же мы с обсуждения возможных путей развития технологий искусственного интеллекта в России.

Как нам обустроить ИИ

Как отмечает ряд ведущих ученых и политиков современности, включая президента РФ В.В. Путина, лидерство в области ИИ в ближайшие годы обеспечит государствам, ускоренно развивающим и внедряющим технологи ИИ, лидерство в мире XXI века, подобно тому, как лидерами второй половины XX века являлись страны, освоившие технологии ядерной физики. В России такая Национальная стратегия развития искусственного интеллекта на период до 2030 г. введена в действие в октябре 2019 г. В ней многое верно сказано о том, что в нашей стране критически необходимо ускоренное внедрение технологических решений, разработанных на основе искусственного интеллекта, в самые разные отрасли экономики и сферу общественных отношений.

Вот только по вопросу о возможных путях и методах такого ускоренного внедрения у отечественных специалистов до сих пор единого мнения нет. Можно даже сказать, что мнения полярно разделились.

Одна группа ученых, бизнесменов, политиков считает, что никакого особого пути у России в данном вопросе быть не может. На их взгляд, очевидно, что ресурсы мирового сообщества разработчиков ИИ намного превосходят наши отечественные. Научная область на сегодня полностью открыта, все результаты публикуются, поэтому никакие отечественные средства и платформы ИИ нам не нужны. Российским разработчикам следует полностью интегрироваться в мировое сообщество, встать на международно признанные аппаратно-программные платформы, созданные гигантами мирового бизнеса, и уже на их основе создавать и выводить на рынок конкурентоспособные в мировом масштабе продукты и услуги в области ИИ. Российскому же государству нужно лишь обеспечить привлекательную бизнес-среду для возникновения и расцвета русских интеллектуальных стартапов, которые вскоре естественным образом достигнут уровня своих зарубежных аналогов. Этим путем предлагают двигаться уважаемые коллеги из «Сбербанка», «Яндекса» и других серьезных и крупных российских компаний – лидеров в публичной сфере обсуждений тематики ИИ.

Другая группа специалистов, связанная, в первую очередь, с решением задач обороны и безопасности страны, а также с созданием крупных промышленных изделий отечественной техники, придерживается иного мнения на этот счет. И нам в ГосНИИАС близка именно эта, вторая точка зрения. Нам представляется, что необходимость срочного и наискорейшего развития в РФ отечественных технологий и средств ИИ определяется следующими основными факторами:

- важностью внедрения технологий ИИ для обеспечения национальной безопасности и обороноспособности страны уже в самой ближайшей перспективе;

- необходимостью обеспечения конкурентоспособности российских изделий (в том числе, военного и двойного назначения, производимых предприятиями ОПК) на мировом рынке;

- отставанием отечественной промышленности в области создания ряда ключевых для развития ИИ программных и аппаратных компонент;

- происходящей непосредственно в настоящий момент революцией искусственного интеллекта, о которой мы так подробно говорили в предыдущих частях данной статьи.

Мы считаем, что Россия не может развивать технологии ИИ так, как их развивают мировые лидеры – США и Китай. Нам нужен асимметричный ответ, иная стратегия, другие инструменты: единая технологическая платформа, национальные банки данных и обученных моделей, совершенно иная структура организации разработок.

Предвижу реакцию уважаемых просвещенных читателей: «Что за глупость! Мы это уже проходили. В мире все в равных условиях. Почему же только нам требуется какой-то особый путь?». И действительно, сложно обосновать потребность в специальных мерах поддержки и организации работ в сфере высоких технологий, например, особым историческим путем или духовными традициями русского народа. Все куда прозаичнее. Причины здесь сугубо рациональные, и связаны они с хорошо известным всем экономистам «эффектом масштаба».

Дело в том, что даже на самом честном и открытом рынке не все действительно находятся в равных условиях. Именно поэтому мелкие продуктовые лавки всегда проиграют конкурентную борьбу «Ашану», а небольшой стартап, вздумавший сегодня производить смартфоны, вряд ли сможет всерьез бороться с Apple или Samsung. Все специалисты, включая и коллег-сторонников открытых платформ, согласны в том, что по ресурсному обеспечению – людьми, финансами, вычислительными мощностями – мы в России отстаем от мировых лидеров в десятки раз. У нас есть замечательные команды высококвалифицированных ученых и разработчиков, имеющих отличную репутацию, в том числе и на международном уровне. Но их крайне мало! Даже если наши вузы увеличат в ближайшие годы выпуск молодых специалистов в 3-4 раза, это не позволит нам сколько-нибудь существенно приблизиться по числу занятых в сфере внедрения технологий ИИ к США и Китаю. Отсюда очевидный прогноз. В области разработки и внедрения ИИ в конечные изделия наши разрозненные и малочисленные разработчики всегда будут выступать как «малый бизнес», который из-за эффекта масштаба проигрывает крупным иностранным разработчикам. В прямой конкуренции с мировым сообществом выживут лишь крупные нишевые игроки – те же «Яндекс», «Сбербанк», «Лаборатория Касперского».

В сфере бизнеса с этим, вероятно, и можно было бы смириться. Но как нам быть со сферой обороны и безопасности? Использовать зарубежные готовые продукты мы здесь заведомо не можем – хотя бы по причине текущих санкций, а также возможных подобных санкций в будущем. При этом количество изделий и систем с ИИ, необходимых для поддержания оборонного паритета с теми же США и Китаем, исчисляется сотнями, если не тысячами. Их создают многочисленные предприятия ОПК в условиях целого ряда ограничений, в том числе экономических. Где же нам взять в ОПК такое количество специалистов и команд разработчиков высокого уровня, вычислительных ресурсов, программных и аппаратных средств, которые обеспечат этот паритет? Возможно ли это в принципе? Нам представляется, что возможно. Для этого необходимы отечественная технологическая платформа и четко скоординированная программа развития отечественных технологий искусственного интеллекта на государственном уровне.

Единственный способ нивелировать проблему масштаба – создание единого национального «конвейера» или «производственной сети» по разработке и внедрению технологий ИИ. Если вся страна станет в области ИИ «единым предприятием», то она окажется одним из крупнейших таких «предприятий» в мире. Даже если речь пойдет о технологическом агрегировании только лишь ресурсов в сфере ИИ в ОПК, все равно эффект масштаба (за счет унификации, разделения труда и многократного использования накопленных данных, знаний и решений) будет очень значительным.

Прежде всего, необходима межведомственная координация работ, связанных с развитием и внедрением технологий ИИ с учетом уровня их технологической готовности. Вы помните, мы говорили о том, что первая волна технологической революции принесла нам решения в области компьютерного зрения, анализа больших данных, обработки и анализа сигналов. Эти технологии можно назвать технологиями «первого эшелона». Сегодня они практически готовы к внедрению в реальные изделия, и это внедрение нельзя откладывать – нужно ставить соответствующие опытно-конструкторские работы (ОКР), планировать серийное производство. Технологии второй волны составляют «второй эшелон» с более низкой технологической готовностью. Здесь требуются поисковые исследования, научно-исследовательские работы, за которыми обязательно уже сейчас нужно планировать соответствующие ОКР. Такая конвейерная организация работ должна стать предметом совместных действий Минпромторга, Минобрнауки, Российской академии наук, научных фондов и институтов развития. Нужно обеспечить реализацию и отработку новых прорывных технологий ИИ, оценку приоритетности и целесообразности внедрения этих технологий в изделия, оценку безопасности и контролируемости новых технологий ИИ.

Не вызывает сомнений и необходимость обеспечения «импортонезависимости». Требуется разработка российских программных продуктов и платформ в области ИИ, создание и внедрение отечественных процессоров и программно-аппаратных комплексов на базе российских компонентов для работы с инструментами ИИ. Также чрезвычайно важен и вопрос создания специализированных банков данных и баз знаний для обучения с единым централизованным доступом к ним. Это касается и суперкомпьютерных ресурсов как основы инфраструктуры для разработки и применения средств ИИ.

Необходимость жесткой стандартизации и унификации технологий ИИ, по крайней мере, тех, которые применяются при решении задач обороны и безопасности, связана с тем самым «эффектом масштаба». Именно таким путем должны быть обеспечены экономия сил и средств в масштабах государства, а также реальная прозрачность выбора, оценки и тестирования создаваемых решений в области ИИ. Это к тому же еще и антикоррупционный механизм, не позволяющий дублировать разработки или выбирать не самые лучшие решения.

С этим же связана и необходимость создания новых методик тестирования и испытания для систем и изделий, использующих машинное обучение. Обучаемые алгоритмы гораздо сложнее «поймать» на ошибках в ходе испытаний.

Например, мы испытываем на полигоне систему технического зрения, распознающую объекты. Стоит ее дообучить на изображения с этого полигона, как она тут же перестанет ошибаться… на этом полигоне. А при применении в других условиях может сработать гораздо хуже. Поэтому обучаемые изделия требуют более разнообразной базы тестирования, широкого применения методов математического и полунатурного моделирования для имитации всех возможных условий. Причем для каждой группы типовых интеллектуальных задач необходимо определить свои характерные объемы тестовых примеров, число испытаний и протоколы тестирования.

Для реализации всех этих подходов потребуется создание новой нормативной базы. Нужно учесть особые правила сертификации систем с ИИ, возможность их дообучения в ходе эксплуатации. Придется также навести порядок в использовании интеллектуальных прав на базы данных, нейросетевые модели и другие результаты интеллектуальной деятельности, возникающие в сфере ИИ.

В результате всех описанных мер должна возникнуть единая российская технологическая платформа в области искусственного интеллекта. Под такой платформой мы понимаем совокупность унифицированных информационных и аппаратно-программных средств, а также корпуса соответствующих отраслевых стандартов и методик, обеспечивающую разработчикам технологий, создателям изделий, их заказчикам и эксплуатантам общий инструментарий, терминологию, единые правила игры, общее информационное пространство. Разработчикам технологий она обеспечит связь с разработчиками изделий, возможность получения данных и проведения экспериментов для развития технологий, а также возможность быстрого продвижения новых технологических блоков в изделия по мере развития технологий. Разработчики изделий также получат двустороннюю связь как с разработчиками технологий, так и с заказчиками – по техническим характеристикам, правилам тестирования и приемки.

Немаловажной для конструкторов будет и возможность легкого и быстрого прототипирования будущих систем со встраиванием новых интеллектуальных блоков. Заказчикам будут обеспечены исключение дублирования работ за счет создания банка готовых решений, а также экономия средств за счет стандартизации и унификации. Возрастет доверие потребителей к новым интеллектуальным изделиям и технологиям за счет полной прозрачности и доступности для контроля характеристик. Кроме того, эксплуатанты получат в рамках этой платформы постоянную поддержку по всему жизненному циклу изделий с элементами ИИ. Для этого технологическая платформа ИИ должна включать целый ряд взаимоувязанных между собой слоев: программный слой, аппаратный, инфраструктурный, информационный, методический, нормативный… Только согласованная работа всех этих инструментов позволит нам достичь того, к чему мы сегодня стремимся – ускоренной разработки, ускоренного внедрения и максимальной эффективности эксплуатации изделий с искусственным интеллектом.

Конечно, создание такой многослойной и всеобъемлющей технологической платформы – пока дело будущего.

Однако некоторые ее элементы уже начинают появляться, например, в части аппаратных и программных средств. И мы надеемся, что прототипом ее программного слоя, технологической основой для разработки и широкого внедрения отечественных интеллектуальных решений сможет стать созданная нами в ГосНИИАС унифицированная отечественная программная платформа нейросетевой разработки Plat.

Экосистема машинного обучения

В конце прошлого года мы завершили разработку первой версии унифицированной программной платформы нейросетевой разработки «Платформа-ГНС» (платформа глубоких нейронных сетей). Работа была поддержана Минпромторгом РФ. В качестве технического заказчика выступило АО «ГосМКБ «Радуга» им. А.Я. Березняка» (входит в состав корпорации КТРВ). Сегодня платформа включает интегрированную среду глубокого обучения нейросетей, отечественную библиотеку глубокого обучения Plat, а также набор средств аппаратной реализации нейросетей для различных аппаратных платформ, в том числе и отечественного производства. Пока она поддерживает лишь технологии «первого эшелона», полностью готовые для практического внедрения, но в дальнейшем мы планируем развить на ее основе целую экосистему машинного обучения и искусственного интеллекта.

Исходной мотивацией для создания отечественной среды обучения нейросетей являлось, конечно, импортозамещение. В настоящее время обучение ГНС осуществляется с использованием программных средств и технологий, разработанных за рубежом. В связи с этим возникают не только риски, связанные с зарубежным контролем критических оборонных технологий, но и непосредственные технологические ограничения, связанные с отсутствием поддержки и учета особенностей отечественных датчиков и бортовых вычислителей. Таким образом, задача замещения их отечественной технологией является не только критической с точки зрения обеспечения импортонезависимости данного ключевого направления, но и полностью технологически оправданной. Кроме того, ввиду широкого разнообразия используемых на сегодняшний день средств и систем разработки решений на базе нейросетей, представляется необходимым создание единой унифицированной отечественной программной платформы, которая, с одной стороны, обеспечила бы возможность импорта передовых разработок из других популярных платформ, а с другой стороны, исключила бы появление дублирующих или несовместимых разработок. При этом созданные с использованием такой единой платформы алгоритмы должны автоматически преобразовываться в бортовое ПО для всех основных типов перспективных отечественных вычислителей.

Другая важная мотивация здесь связана с необходимостью учитывать особую специфику уязвимости глубоких нейронных сетей. В последние годы одной из главных проблем, стоящих на пути промышленного внедрения ИИ на основе машинного обучения, стало существование т.н. «атак» на нейронные сети. Возможность таких атак определятся тем, что, в отличие от обычного ПО, содержательные алгоритмы обработки данных в ГНС представляют собой не инструкции на каком-либо языке программирования, а наборы обученных коэффициентов (большие массивы чисел). Поэтому несанкционированные изменения в структуре ГНС не могут быть выявлены стандартными средствами контроля и анализа ПО, которые сейчас применяются. При этом сами нежелательные изменения в структуре и поведении ГНС (атаки) могут осуществляться различными способами, включая, например, манипуляции с обучающей выборкой либо процессом обучения, что также невозможно контролировать стандартными средствами и подходами. В силу этого гарантировать то, что обученная ГНС будет работать именно так, как задумали разработчики, можно лишь при полном контроле как программной, так и аппаратной составляющих процесса их формирования. Поэтому использование ПО с элементами ИИ (ГНС) в критических приложениях, связанных с рисками и безопасностью, требует, на наш взгляд, обязательного использования на этапе разработки полностью отечественных аппаратно-программных комплексов.

Был и еще один важный фактор, определяющий необходимость создания не просто отечественной библиотеки обучения нейросетей, но именно целой экосистемы поддержки отечественных нейросетевых разработок. Как уже говорилось выше, в отличие от США, Китая, Японии, ведущих европейских стран, мы еще долгие годы (пока не воспитаем на порядок большее, чем сейчас, число специалистов) не сможем позволить себе собирать на каждом конечном предприятии полноценную команду разработчиков мирового уровня для создания каждого необходимого интеллектуального продукта. Поэтому мы с самого начала создавали нашу платформу как интегрированную среду поддержки разработок в области ИИ, которая позволит ускорить внедрение глубоких нейронных сетей в России путем преодоления ряда технологических барьеров и облегчения ряда технологических переходов для всех участников отрасли. Наша цель – не просто снять ограничения для применения ГНС для предприятий ОПК. Предлагаемая целостная экосистема должна содержать банк готовых типовых решений с аппаратными реализациями, что позволит обеспечить наискорейшее внедрение самых передовых научных результатов в конечные изделия и производственные процессы. Более того, за счет использования встроенных схем обучения и средств визуального программирования мы надеемся понизить порог входа для обычных инженеров и тем самым резко увеличить сообщество разработчиков прикладных систем ИИ. Теперь для этого не потребуется глубоко понимать принципы машинного обучения и даже уметь программировать. Грубо говоря, идея заключается в том, чтобы вместо малого числа «ресторанов высокой ИИ-кухни», которыми Россия располагает сегодня, создать для нашей промышленности широкую «сеть быстрого питания» современными интеллектуальными решениями, где