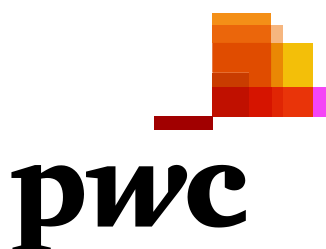


Искусственный интеллект может привести к серьезным изменениям текущих бизнес-моделей и к очередной промышленной революции. Однако совет директоров вряд ли даст зеленый свет решениям на базе искусственного интеллекта, пока не убедится в эффективности стратегии в этой области, возможности ее реализации и получении ожидаемых результатов. Без такой предварительной проработки директора рискуют оказаться в ситуации, когда им придется действовать вслепую. Так каким же образом ваша компания может обеспечить необходимый контроль, чтобы полностью реализовать потенциал искусственного интеллекта?

Ускорение инновационного развития

Уверенность в реализации возможностей искусственного интеллекта



Введение

Раскрываем потенциал

Сегодня человек и машина взаимодействуют все активнее. Инновационные разработки выходят за пределы научно-исследовательских лабораторий и становятся общедоступными. Технологии искусственного интеллекта открывают возможности для инновационных преобразований в интересах потребителей, бизнеса и общества в целом.

Однако для реализации этих возможностей необходима более эффективная система управления и новая операционная модель.

При помощи искусственного интеллекта можно решить многие актуальные проблемы современного общества, начиная от производства экологически чистой энергии и заканчивая лечением рака и других заболеваний¹.

Основной потенциал технологий искусственного интеллекта для бизнеса заключается в возможности преобразовывать данные в объекты интеллектуальной собственности. По данным проведенного PwC опроса², 70 % руководителей компаний считают, что в будущем применение искусственного интеллекта станет конкурентным преимуществом.

Неопределенность мешает внедрению инноваций

Актуальный вопрос – как реализовать весь этот потенциал? Компания Netflix служит классическим примером использования накопленных данных для создания преимуществ за счет машинного обучения. Благодаря прорывным технологиям изменился подход к предоставлению медиаконтента и были полностью пересмотрены бизнес-модели в секторе развлечений.

Однако в мировой экономике темпы освоения технологий искусственного интеллекта разнятся. Большое число компаний по-прежнему находятся лишь в начале этого пути. Так, менее 40 % респондентов, принявших участие в нашем последнем Ежегодном опросе руководителей крупнейших компаний мира, начали анализировать, как искусственный интеллект отразится, например, на их будущих потребностях в специалистах, обладающих определенными знаниями и навыками³.

¹ Исследование PwC «Bot.me: революционное партнерство. Как искусственный интеллект сближает человека и машину» (Bot.me: A revolutionary partnership. How AI is pushing man and machine closer together), апрель 2017 г. (<https://www.pwc.com/us/en/press-releases/assets/img/bot-me.pdf>).

² Глобальное исследование PwC Digital IQ (www.pwc.ru/digital-iq).

Некоторые компании рискуют оказаться в затруднительном положении со стратегической точки зрения. Отчасти это связано с нелегким выбором: разнообразие технологий, инновационных решений и поставщиков просто поражает. Сложность также заключается в оценке, минимизации и контроле репутационных и технологических рисков, связанных с новыми и по большей мере неиспытанными технологиями. К основным трудностям относится валидация данных и определение мер безопасности, обеспечивающих надлежащее выполнение машинами распоряжений человека.

Также возникают вопросы этического характера, в том числе: «Приемлемо ли влиять на выбор человека?» и «Имеют ли потребители достаточное представление о том, как используются их данные и кто имеет к ним доступ?».

В своем стремлении активно продвигать искусственный интеллект компании зачастую пытаются реализовать слишком много возможностей или не могут адекватно оценить экономическое обоснование проекта и сопряженные с ним риски. В результате большое количество разработок так и остаются незавершенными.

Тщательная оценка и продуманная реализация

Эти трудности подчеркивают необходимость построения новой модели стратегической оценки, управления и реализации. Без нее технологии искусственного интеллекта могут так и остаться в стенах корпоративных лабораторий или же компании столкнутся с неприемлемыми рисками, которые могут нанести серьезный урон.

74 %

российских компаний планируют инвестировать в искусственный интеллект в ближайшие три года².

67 %

руководителей компаний полагают, что в ближайшие пять лет искусственный интеллект и автоматизация (включая блокчейн) будут отрицательно сказываться на доверии заинтересованных сторон к отрасли⁴.

В основе такого подхода лежит потребность в доверии и прозрачности при создании ответственного искусственного интеллекта. Многие заинтересованные стороны как в самой организации, так и из числа ее клиентов, органов регулирования и других сторонних участников могут скептически относиться к использованию искусственного интеллекта. Поэтому очень важно продумать, как добиться повышения уровня доверия всех заинтересованных сторон к этим технологиям. Главное – повысить прозрачность и осведомленность в вопросах о том, как используются системы на базе искусственного интеллекта, какие задачи они выполняют, какие решения принимают и какие возможности создают. Мы считаем, что в этом и заключается суть «ответственного искусственного интеллекта».

Хотя никто не может гарантировать, что сложные автономные программы-агенты будут безупречно выполнять свои функции, существует целый ряд передовых методов, включая разработку средств контроля и их мониторинг, которые помогут минимизировать риск и будут способствовать ответственному применению технологий искусственного интеллекта.

В данной публикации мы рассмотрим трудности, которые необходимо будет преодолеть при внедрении такого подхода, а также связанные с ним возможности и варианты его использования. Цель заключается не в сдерживании или замедлении инноваций, а, напротив, в их ускорении за счет укрепления уверенности советов директоров и предоставления им платформы для достижения желаемых результатов.

3, 4 «20 лет аналитической оценки взглядов и мнений руководителей компаний...Что дальше?» (20 years inside the mind of a CEO...What's next?), 20-й Ежегодный опрос руководителей крупнейших компаний мира, PwC, 2017 г. (ceosurvey.pwc).

Управление прорывными технологиями

Как искусственный интеллект меняет правила игры

Внедрение искусственного интеллекта влечет за собой серьезные последствия для всех участников процесса управления бизнесом.

Развитие технологий на базе искусственного интеллекта создает условия для построения совершенно новых операционных и бизнес-моделей. Компании получают возможность анализировать данные на уровнях за пределами человеческого понимания и действовать исходя из каждого нового полученного набора данных. А это, в свою очередь, позволяет компаниям применять индивидуальный подход в работе с клиентами, разрабатывать товары и услуги с учетом их потребностей, а также выявлять возможности роста так быстро и точно, как никогда ранее.

Что такое искусственный интеллект?

В своей книге «Искусственный интеллект: современный подход» (Artificial Intelligence: A Modern Approach) Стюарт Рассел и Питер Норвиг дают следующее определение искусственному интеллекту: «проектирование и построение интеллектуальных агентов, которые воспринимают объекты окружающей среды и предпринимают действия, влияющие на окружающую среду»⁵. Основное различие между искусственным интеллектом и программным обеспечением общего назначения заключается в словосочетании «предпринимают действия». Благодаря искусственному интеллекту машины могут самостоятельно реагировать на сигналы из окружающего мира, то есть сигналы, которые программисты напрямую не контролируют и, следовательно, не могут предвосхитить. Из всех областей искусственного интеллекта стремительнее всего развиваются методы машинного обучения, то есть способность программного обеспечения совершенствоваться в процессе взаимодействия с миром. Технологии искусственного интеллекта можно разбить на три группы:

- **Вспомогательные** (Assisted Intelligence): широко распространенные технологии, которые совершенствуют процессы, уже выполняемые людьми и компаниями.
- **Расширенные** (Augmented Intelligence): новые технологии, которые позволяют людям делать то, что иным образом сделать нельзя.
- **Автономные** (Autonomous Intelligence): технологии, разрабатываемые на будущее и предусматривающие создание машин, которые могут действовать самостоятельно.

Некоторые примеры успешного применения искусственного интеллекта:

- **Информирование:** «персональные помощники», например Alexa и Siri, а также чат-боты банков и операторов мобильной связи.
- **Прогнозирование поведения:** Национальная служба здравоохранения Великобритании внедряет пилотный проект по машинному обучению с целью прогнозирования неявки амбулаторных пациентов в британские больницы. Это позволит оптимизировать процесс создания и переноса записи к врачам, получить представление о факторах, влияющих на поведение пациента, и создать условия для принятия мер по снижению уровня неявки.
- **Охрана здоровья:** искусственный интеллект используется как вспомогательное средство диагностики в медицине. Согласно нашему исследованию, существенная доля людей по всему миру готовы проходить определенные виды лечения, сдавать анализы и получать услуги под управлением искусственного интеллекта или робота⁶.
- **Повышение лояльности:** телекоммуникационные и медийные компании проводят анализ данных о клиентах на базе машинного обучения для прогнозирования и рекомендации мер, направленных на предотвращение оттока клиентов.
- **Прогнозирование спроса:** предприятия розничной торговли начинают использовать технологии глубинного обучения для прогнозирования клиентских заказов.

- **Всеобщая кастомизация:** благодаря роботам-консультантам появилась возможность предлагать индивидуальные инвестиционные решения более широкому кругу клиентов. До недавнего времени получение консультаций по вопросам инвестирования на таком уровне было доступно только состоятельным клиентам.
- **Повышение качества:** производители используют технологии искусственного интеллекта для улучшения контроля за качеством, сокращения производственных простоев, а также повышения скорости и эффективности производственных процессов.
- **Интеллектуальные процессы:** интеллектуальная автоматизация процессов приводит к существенной экономии затрат, связанных с финансовыми и кадровыми вопросами, а также соблюдением нормативно-правовых требований. Роботизация бизнес-процессов в сочетании с технологиями искусственного интеллекта позволяет выполнять массовые типовые задачи.

Сориентироваться во всем многообразии алгоритмов и приложений, подпадающих под категорию «искусственный интеллект», – само по себе невероятно трудная задача. До настоящего времени много внимания уделялось автоматизации уже выполняемых задач⁷. Однако, по мере того как работники будут освобождаться от выполнения типовых задач, а сотрудничество между человеком и машиной будет становиться все более тесным, появятся возможности для новых прорывов. Они будут реализованы благодаря принятию более продуманных решений и появлению совершенно новых бизнес-моделей на базе расширенных технологий искусственного интеллекта. Наглядным примером отрасли, в которой прорывные технологии уже привели к коренным изменениям, служит индустрия развлечений. Существует множество вариантов использования искусственного интеллекта, изменяющих нашу повседневную жизнь и связанную с этим коммерческую деятельность. Беспилотные автомобили – лишь один из них.

За последние несколько лет коммерческое применение искусственного интеллекта получило широкое распространение. Этому способствует наличие больших вычислительных мощностей, доступность супермассивов данных и развитие машинного обучения (включая глубинное обучение). Методы машинного обучения часто используются для прогнозной аналитики и классификации изображений и речи. Однако их также можно сочетать с элементами «традиционного» искусственного интеллекта, такими как обработка естественного языка, стратегическое планирование и логическое рассуждение, для создания мощных автономных агентов.

Так насколько широко распространились технологии искусственного интеллекта? Если не считать крупные технологические компании, которые уже несколько лет используют искусственный интеллект для оказания услуг, большинство инновационных решений находятся на стадии становления и представляют собой лабораторные опытные образцы или предмет научных исследований. Современному бизнесу следует сосредоточиться на создании условий для успешного переноса этих разработок в реальные цепочки создания стоимости.

5 Стюарт Рассел, Питер Норвиг. Искусственный интеллект: современный подход (Artificial Intelligence: A Modern Approach). Издательство Pearson, 2009 г.





6 «Доктор? Какой доктор? Или как искусственный интеллект и робототехника будут лечить вас по-новому» (What doctor: Why AI and robotics will define the New Health) (<https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health/survey-results.html>).

7 Влияние автоматизации и развития искусственного интеллекта на производство и занятость рассмотрено в нашей публикации «Займут ли роботы наши рабочие места? Возможные последствия автоматизации в Великобритании и других крупнейших странах мира» (Will robots steal our jobs? The potential impact of automation on the UK and other major economies), март 2017 г. (<https://www.pwc.co.uk/economic-services/ukeo/pwcukeyo-section-4-automation-march-2017-v2.pdf>).

Новый подход

Как показано на рис. 1, для внедрения систем искусственного интеллекта компаниям необходимо по-новому взглянуть на технологии, развитие бизнеса и реализацию стратегии. Кроме того, нужно пересмотреть соответствующую операционную модель и процессы принятия решений. Такие изменения затронут не просто отдельные подразделения, занимающиеся технологиями и инновациями, а все предприятие в целом.

Рисунок 1. Изменение подхода к ведению бизнеса

	Традиционный подход	Новый подход
Стратегия 	Технологии для управления информацией Данные – инструмент бизнес-аналитики Детерминистский подход	Технологии для управления вашим бизнесом Данные – объект вашей интеллектуальной собственности, обеспечивающий конкурентное преимущество Направляющий (итеративный) подход
Проектирование 	Опыт пользователя рассматривается как один из уровней приложения Процедура принятия решений жестко запрограммирована Извлекаемая информация – факт из базы данных	Опыт пользователя – главная характеристика приложения Программное обеспечение обучается процессу принятия решений Извлекаемая информация – наиболее вероятный правильный ответ
Разработка 	Линейная разработка технологий Руководители бизнес-подразделений предоставляют спецификацию, ИТ-отдел занимается разработкой	Итеративная разработка технологий и бизнес-моделей Профильные эксперты из бизнес-подразделений входят в состав команды разработчиков
Операционная модель 	Постоянная, периодически обновляемая технология Технические риски заключаются в основном в простых системы и системных ошибках Кибератаки	Динамичные, гибкие модели. Непрерывное развитие с учетом результатов тестирования Технические риски включают в себя приобретенные в результате обучения и неожиданные модели поведения «Состязательные атаки»

Источник: PwC



Что следует учитывать при построении операционной и бизнес-модели на базе искусственного интеллекта?

Стратегия

1. Соответствие стратегическим целям

Крайне важно, чтобы инновационные разработки в сфере искусственного интеллекта соответствовали основным стратегическим целям и показателям эффективности. Не следует допускать большого числа разрозненных инициатив, каждая из которых реализуется обособленно от других. Мы видим, что многие компании запускают различные пилотные проекты. Однако при этом большинство компаний не имеют полного представления о возможных дестабилизирующих последствиях внедрения искусственного интеллекта и не выявляют связанные с этим угрозы и возможности.

2. Не стоит ждать чудес

Искусственный интеллект – пусть и разумная, но все же машина. Бытует ложное мнение о том, что искусственный интеллект способен чудесным образом обучаться без вмешательства человека. На самом деле вам придется приложить немало усилий для получения и очистки данных, составления классификаций и обучения машин и сотрудников⁸.

3. Выбираем и оцениваем партнеров

Сегодня существует огромное количество стартапов, которые предлагают самые разные решения и возможности. Партнерство с такими компаниями способствует инновационному развитию, повышает гибкость и сокращает сроки выхода на рынок. Однако важно четко понимать, какие цели вы преследуете, вступая в партнерство. Так, при выборе партнера нужно определить свои стратегические и операционные приоритеты. Важно также учитывать, что в некоторых случаях поставщики могут отличить рекламировать возможности, но не совсем понимать, как их реализовать: их взгляды на риски, связанные с разработкой решений,

Искусственный интеллект – пусть и разумная, но все же машина. Бытует ложное мнение о том, что искусственный интеллект способен чудесным образом обучаться без вмешательства человека.

безусловно, сильно отличаются от привычного вам подхода.

В условиях повышенных рисков и стремительных перемен, происходящих на рынке поставщиков, следует прежде всего оценивать финансовую устойчивость потенциальных партнеров: смогут ли они поддержать вас, когда вам это понадобится? Не менее важно определить, каким образом вы будете получать нужные данные, приобретать знания, необходимые для реализации новых возможностей, и интегрировать новые платформы в имеющуюся инфраструктуру. При покупке доступных на рынке готовых программных продуктов зачастую необходимо пройти этап экспериментальной проверки.

4. Создаем условия для тщательной проверки

Прежде чем внедрять искусственный интеллект, необходимо четко понимать, как он работает и какие задачи выполняет. В том числе следует убедиться, что программа предусматривает понятную процедуру принятия решений, которую специалисты бизнес-подразделений могут проверить. Важно продумать, как будет обеспечено достижение желаемых результатов работы ПО, особенно в случае машинного обучения. Чтобы начать действовать, совету директоров необходима уверенность в результатах. Этого, по всей вероятности, ожидают и регуляторы. Для того чтобы получить такую уверенность, нужно, в частности, обеспечить прозрачность применяемых подходов и процедур. Однако для выполнения этой задачи, вероятно, придется искать компромисс между прозрачностью процесса принятия решений, производительностью системы и функциональными возможностями.

5. Демонстрируем нормативно-правовое соответствие

Чтобы поспевать за развитием технологий, регуляторам приходится действовать быстро. В жестко регулируемых отраслях экономики, таких как здравоохранение и финансовые услуги, внедрение искусственного интеллекта может быть затруднено ввиду нормативных ограничений. Ситуация усугубляется в связи с изменениями в регулировании, например выходом в ЕС Общего регламента о персональных данных (General Data Protection Regulation, GDPR). Неизменное соблюдение применимых норм регулирования лежит в основе доверия к платформам на базе искусственного интеллекта.

6. Организационная структура

Изменение бизнес-модели в рамках реализации общей стратегии в области искусственного интеллекта также должно отразиться и на организационной структуре вашей компании. Вам нужно создать отдельное подразделение, которое будет заниматься управлением технологиями искусственного интеллекта. В такое подразделение может войти специально назначенный представитель высшего руководства и все сотрудники центра передового технического опыта. Наличие в вашей компании специалистов по обработке и анализу данных (обученных в рамках организации или привлеченных со стороны) крайне важно для формирования развитой организационной структуры, готовой к применению искусственного интеллекта.

Проектирование

1. Вскрываем «черный ящик»

Приложения на базе искусственного интеллекта могут взаимодействовать с клиентами и принимать важные бизнес-решения. Однако основная часть этих процессов происходит внутри «черного ящика». Такое отсутствие прозрачности неизбежно приводит к возникновению репутационных и финансовых рисков. При проектировании программного обеспечения важно обеспечить его максимально возможную прозрачность и контролируемость.

Надлежащие механизмы управления и защиты предусматривают возможность следить за состоянием модульных систем, а также оперативно выявлять, корректировать и (при невозможности корректировки) отключать вышедшие из-под контроля модули без необходимости прекращения работы всей платформы. При этом приоритет отдается выявлению факторов зависимости и способности вносить корректировки с минимальным ущербом для обычного порядка ведения деятельности в случае изменения нормативно-правовой базы и других аспектов деятельности.

2. Создаем интересный опыт пользователя

Во многих приложениях на базе искусственного интеллекта предусмотрены очень субъективные показатели эффективности взаимодействия с пользователями, которые можно сравнить с IQ, индивидуальными особенностями и предсказуемостью. Хотя при разработке основное внимание может уделяться анализу данных, успех продукта будет зависеть от того, какую эмоциональную реакцию он будет вызывать. Такая субъективность указывает на необходимость частой обратной связи между владельцем продукта и разработчиками для эффективного управления изменяющимися ожиданиями и функциональными возможностями. Нередко имеет смысл привлекать в отдел по анализу данных поставщиков, специализирующихся на разработке пользовательских интерфейсов, или внутренних специалистов по цифровым решениям.

Самые эффективные средства контроля закладываются на этапе проектирования и внедрения и позволяют обнаруживать проблемные вопросы до того, как они превратятся в серьезные проблемы. Такие средства контроля также помогают выявлять возможности для оптимизации.

Искусственный интеллект зачастую может превосходить людей в выполнении определенных задач или в конкретных предметных областях, но в целом он не способен применять эти навыки и знания для решения других проблем. Для людей, которым приходится взаимодействовать с искусственным интеллектом (особенно впервые), это не очевидно, поэтому при таком взаимодействии их может ждать недоумение и разочарование.

В связи с этим при проектировании следует обязательно уделять внимание разработке бренда и созданию персонажа («конструирование функционала»). Если вы все сделаете правильно, то даже самое базовое ПО будет казаться человеком. Если же вы ошибетесь, то пользователи уйдут от вас.

В некоторой степени анализ, выполняемый искусственным интеллектом, неизбежно является вероятностным и основывается на неполных сведениях. Поэтому важно, чтобы вы понимали имеющиеся ограничения и могли разъяснить их своим клиентам. Это следует учитывать, например, при предоставлении инвестиционных рекомендаций, подготовленных роботами-консультантами.

3. Внедряем систему контроля

Самые эффективные средства контроля закладываются на этапе проектирования и внедрения и позволяют обнаруживать проблемные вопросы до того, как они превратятся в серьезные проблемы. Такие средства контроля также помогают выявлять возможности для оптимизации.

Важный вопрос – кто проектирует средства контроля и следит за их работой? Поскольку приложения имеют широкую сферу применения, а также существует необходимость отслеживать результаты, для выполнения поставленных задач следует привлекать специалистов

из различных подразделений организации. Для проектирования средств контроля необходим большой объем информации со стороны бизнес-экспертов. Для физических приложений может потребоваться поддержка специалистов по технике безопасности.

Ключевой момент при внедрении – разделение средств контроля на два уровня («иерархический подход»).

Как минимум необходим уровень жесткого контроля, на котором устанавливаются ограничения и прописываются действия, которые следует предпринять в случае нарушения этих ограничений. Примером может служить максимальная сумма транзакции для алгоритма торговых операций на финансовом рынке. В более сложные приложения, такие как диалоговый агент, можно встраивать «стабилизатор поведения», который корректирует основной алгоритм в случае возникновения риска ошибки, например нарушения нормативно-правовых требований или использования ненормативной лексики.

Эти основные средства контроля можно дополнять «моделями критической оценки». Такие модели используются в качестве исходной точки для мониторинга пригодности и точности методов искусственного интеллекта или для отслеживания нежелательных ошибок или отклонений в процессе обучения моделей на базе новых данных. Кроме того, такой подход можно сочетать с принципом непрерывного развития для оптимизации существующих моделей или выявления лучших моделей в целях обновления систем.

8 Публикация доктора Ананда Рао (PwC) «Пять мифов и фактов об искусственном интеллекте» в бюллетене «Прогнозная аналитика и футуризм», выпускаемом организации Society of Actuaries, декабрь 2016 г.

Разработка

1. Пересматриваем подход к управлению программой

Применение традиционного подхода к планированию, проектированию и построению решений, основанных на данных, обречено на неудачу. Чтобы справиться с высоким уровнем сложности возникающих проблем, необходима многоэтапная проработка инноваций и подтверждение концепций в рамках итеративного подхода. Это требует глубокой заинтересованности и вовлеченности владельцев продуктов.

2. Управляем зависимостью от данных

Функциональные возможности искусственного интеллекта сильно зависят от данных для машинного обучения и, как правило, требуют наличия хранилища информации – так называемой базы знаний. Зачастую это означает, что первоначальные проектировочные спецификации и ожидания выходят за рамки того, что можно получить на основе имеющихся данных, независимо от «уровня развития интеллекта» ПО. Ключевое требование к проектам, зависящим от данных, – определение объема и качества данных, а также связанных с этим ограничений, касающихся итоговых моделей и функционала. Это одна из причин, по которым внедрение ПО на базе искусственного интеллекта сопряжено с большим числом итераций при проектировании на этапе разработки решений.

3. Выделяем время на тестирование и обучение

Разработчикам следует применять передовые методы настройки и перекрестной проверки во избежание переобучения («переподгонки») и других распространенных проблем. Особенно это касается машинного обучения. Для получения полного представления о вариантах использования разрабатываемых решений и об опыте пользователя программисты должны получать информацию не только от проектировщиков ПО, которые неизбежно считают проект своим и поэтому могут быть недостаточно объективными. Мониторинг должен включать в себя тестирование, которое проводится для корректировки функциональных недочетов.

Один из способов расширить тестирование и снизить риски – пилотное внедрение новых приложений на базе искусственного интеллекта сначала в небольших масштабах, а затем проведение тщательной оценки таких приложений аналитиками, а также пользователями, не имеющими технической подготовки, в рамках обычного хода деятельности. Экспертные суждения и дополнительная контекстная информация используются для дальнейшей валидации, оценки воздействия и настройки системы, прежде чем она будет развернута в более широких масштабах.

Во многих отношениях разработка программного обеспечения на базе искусственного интеллекта больше похожа на разработку видеоигр, чем на автоматизацию и веб-программирование, особенно если такое ПО взаимодействует непосредственно с человеком. Тестирование должно отражать эту особенность: перед релизом следует проводить тщательную оценку опыта пользователя и поэтапное «бета-тестирование» на невидимой аудитории.

Наконец, разработка нужного решения на базе искусственного интеллекта может потребовать нескольких попыток, прежде чем будет достигнут желаемый результат. Поэтому важно убедиться в том, что вы обеспечили уверенность в эффективности ПО и внедрили надлежащие средства контроля, которые могут заблаговременно сообщить о несоответствии данных, технологий или методов обучения, заложенных в модели, экономическому обоснованию.

4. Устанавливаем пороги доверия

Необходимо найти баланс между автоматизацией и выполняемой человеком валидацией (верификацией). Однако установление правильных порогов и уровней доверия для вмешательства человека может оказаться сложной задачей. Если проявить излишнюю осторожность, то польза искусственного интеллекта будет ограниченной. Если слишком ослабить контроль, то искусственный интеллект примет на себя больше рисков, чем сможет сдержать. Для подтверждения соответствия технологии заявленным параметрам необходим непрерывный мониторинг эффективности бизнеса.

Диалоговые агенты на базе искусственного интеллекта участвуют в субъективном общении с людьми, поэтому важно обеспечить установление порогов доверия для соблюдения общественных норм и соответствия ожиданиям пользователей.

Для получения полного представления о вариантах использования разрабатываемых решений и об опыте пользователя необходимо получать информацию не только от проектировщиков программного обеспечения...

Искусственный интеллект в действии

1. Устраняем случайные ошибки (факты необъективности)

По мере разрастания объемов доступной информации и развития вашей модели необходимо следить за появлением непреднамеренных ошибок в отношении определенных групп. Для выявления случаев такой необъективности крайне важна прозрачность.

Для систем, обучающихся на базе взаимодействия с клиентами, рекомендуется периодически проводить функциональный мониторинг (возможно, на основе стандартизированных моделей взаимодействия). Это позволит выявлять любые негативные «заносы», которые случаются в процессе обучения.

2. Защищаемся от атак

Модели машинного обучения (особенно глубинного обучения) можно обмануть при помощи вредоносных действий, известных как «состязательные атаки». Существует возможность найти такие комбинации вводных данных, которые могут привести к получению неправильных результатов в моделях машинного обучения, то есть по сути «взломать» эти модели. Этот риск можно снизить при помощи моделирования «состязательных атак» на ваши модели и переобучения моделей таким образом, чтобы они могли распознавать подобные атаки. Можно разработать специальное ПО для защиты ваших моделей от таких атак. Все это следует учитывать на этапе проектирования.

Модели машинного обучения (особенно глубинного обучения) можно обмануть при помощи вредоносных действий, известных как «состязательные атаки».

3. Признаем значение данных как вашего главного объекта интеллектуальной собственности

Эффективность искусственного интеллекта зависит от данных, на основе которых происходит его обучение. Поддержание высокого качества данных и непрерывная оценка эффективности модели станут залогом успеха при внедрении платформы на базе искусственного интеллекта. По мере перевода данных и технологий на облачные ресурсы коммерческая выгода будет определяться размером и масштабами вашей «интеллектуальной собственности».

Партнерство с поставщиком может предусматривать неизбежный обмен данными, то есть преднамеренную или непреднамеренную передачу ценных объектов интеллектуальной собственности. В связи с этим важно понимать ценность предоставляемых вами данных и тщательно контролировать их передачу и использование.

4. Следим за системными рисками

Мгновенный обвал финансовых рынков в 2010 году наглядно показал, что может случиться при ненадлежащем взаимодействии множества платформ на базе искусственного интеллекта и недостаточности контроля за их работой. Меры безопасности должны включать в себя сценарное планирование, получение представления об уязвимости своих систем и разработку оперативных мер реагирования.

Искусственный интеллект под контролем

Управление искусственным интеллектом требует такой же дисциплинированности, что и управление любыми проектами трансформации, в основе которых лежит использование инновационных технологий. Может ли ваша компания гарантировать, что искусственный интеллект находится под контролем? С чего следует начать, чтобы обеспечить уверенность в искусственном интеллекте, его контролируемость и ответственность? Это бесконечный процесс, однако мы считаем, что существует пять основополагающих шагов, без реализации которых нельзя приступать к внедрению искусственного интеллекта.

1. Разработать четкую стратегию внедрения искусственного интеллекта

Несмотря на потребность в постоянной оценке и адаптации, крайне важно четко понимать, в каком направлении вы движетесь.

- Готовы ли вы к коренным преобразованиям в результате внедрения искусственного интеллекта?
- Учли ли вы социально-этические аспекты такого перехода?
- Каких результатов вы хотите достичь с точки зрения целей бизнеса (например, кастомизация продукта или повышение эффективности бэк-офиса)?

2. Обеспечить прозрачность

Внедрение искусственного интеллекта будет предметом эмоциональных обсуждений как внутри вашей компании, так и среди ваших клиентов и в обществе в целом. Поэтому важно продумать, как обеспечить доверие заинтересованных сторон к внедряемому решению.

Средства контроля необходимо встроить в решение на самом раннем этапе, а не заниматься их проектированием и внедрением тогда, когда система уже разработана и функционирует. Это касается в том числе механизма контроля за результатами проекта и соблюдением установленных требований.

3. Заранее разработать организационную структуру для работы с искусственным интеллектом

Существует большое число возможных моделей для развертывания ресурсов по работе с искусственным интеллектом в рамках всей организации, начиная от создания центра передового опыта и назначения специального члена совета директоров и заканчивая внедрением стратегии «разрабатывай и управляй».

При любом подходе необходимо наладить информационное взаимодействие и сотрудничество между различными подразделениями организации и централизованное координирование инициатив в области искусственного интеллекта.

4. Встроить функцию управления данными в систему искусственного интеллекта

Если данные представляют собой новый объект интеллектуальной собственности, вам следует внедрить механизмы сбора, очистки и контроля вводных данных, а также обеспечить интеграцию функций управления данными и искусственным интеллектом.

5. Интегрировать механизмы обеспечения уверенности в операционную модель искусственного интеллекта

Уверенность в искусственном интеллекте – не разовая задача. По мере развития платформы на базе искусственного интеллекта необходимо оценивать возникающие риски и возможности.

Уверенность в искусственном интеллекте выходит за рамки простого внедрения новой технологии в операционные процессы. Она требует проведения анализа в рамках всей компании для оценки полученных результатов, выявления возникающих рисков и отслеживания открывающихся возможностей.

Опираясь на нашу обширную исследовательскую деятельность и богатый опыт работы с клиентами, мы разработали систему ответственного искусственного интеллекта (см. рис. 2). Эта система призвана обеспечить прозрачность в отношении целесообразности проектов по внедрению искусственного интеллекта и уверенность в наличии средств контроля, гарантирующих достижение ожидаемых результатов для бизнеса.

Рисунок 2. Система ответственного искусственного интеллекта, разработанная PwC



Заключение

Дальнейшие действия

Если информация – это сила, то искусственный интеллект – это ее «зенит», или наивысшая точка. Но, как и любую другую силу, ее нужно применять обдуманно, осторожно и со всей ответственностью.

Искусственный интеллект может стать конструктивной силой, которая поможет людям достигнуть большего и решить многие проблемы современности. Однако существует большой риск того, что искусственный интеллект сможет выйти за пределы разумного контроля. Для вас, как бизнеса, такое развитие событий неприемлемо ни с этической, ни с репутационной точки зрения. Это также приведет к тому, что совет директоров поставит инновации под сомнение и будет откладывать или даже тормозить их внедрение.

Мы в PwC осознаем, насколько важны уверенность и контроль при реализации возможностей искусственного интеллекта. Мы разработали нашу систему ответственного искусственного интеллекта, чтобы укрепить уверенность клиентов в эффективных способах внедрения решений на базе искусственного интеллекта и повысить уровень доверия к результатам использования таких решений. При этом нами движет не настороженность в отношении инноваций, а, напротив, энтузиазм. Мы хотим, чтобы потребители, бизнес и общество могли извлечь максимальную пользу из имеющихся возможностей. Однако для этого необходимо иметь четкое представление об ограничениях, уязвимых местах и подводных камнях искусственного интеллекта и принимать обоснованные управленческие меры по снижению соответствующих рисков, а также контролю за ними.

Наша система ответственного искусственного интеллекта представляет собой действенный механизм, объединяющий все эти приоритеты и призванный обеспечить эффективный мониторинг результатов внедрения решений на базе искусственного интеллекта и управление этими решениями. Мы считаем, что, следуя этим основополагающим принципам, вы сможете ускорить инновационное развитие своей компании и реализовать концепцию искусственного интеллекта.

Для получения более подробной информации свяжитесь с нами

Тим Клау

Руководитель направления по развитию технологий PwC в России

+7 (495) 967 60 18

Адрес электронной почты:

tim.clough@ru.pwc.com

Олег Данильченко

Руководитель Центра компетенции по прикладному анализу данных

+7 (495) 967 60 94

Адрес электронной почты:

oleg.danilchenko@ru.pwc.com

Сергей Прусаков

Директор, технологические услуги

+7 (495) 967 6042

Адрес электронной почты:

sergey.prusakov@ru.pwc.com

Настоящая публикация подготовлена исключительно для создания общего представления об обсуждаемом в ней предмете и не является профессиональной консультацией. Информация, содержащаяся в данной публикации, не может служить основанием для каких-либо действий в отсутствие профессиональных консультаций специалистов. Каких-либо подтверждений или гарантий (явных или подразумеваемых) в отношении точности или полноты информации, содержащейся в данной публикации, не дается. Если иное не предусмотрено законодательством, PwC снимает с себя всякую материальную и иную ответственность и отказывается от каких-либо обязательств в связи с использованием вами или любым другим лицом информации, содержащейся в настоящей публикации, либо отказом от ее использования, а также в связи с любыми решениями, принятыми на основании этой информации.

PwC в России (www.pwc.ru) предоставляет услуги в области аудита и бизнес-консультирования, а также налоговые и юридические услуги компаниям разных отраслей. В офисах PwC в Москве, Санкт-Петербурге, Екатеринбурге, Казани, Новосибирске, Ростове-на-Дону, Краснодаре, Воронеже, Владикавказе и Уфе работают более 2 500 специалистов. Мы используем свои знания, богатый опыт и творческий подход для разработки практических советов и решений, открывающих новые перспективы для бизнеса.

Под «PwC» понимается сеть PwC и/или одна или несколько фирм, входящих в нее, каждая из которых является самостоятельным юридическим лицом. Глобальная сеть PwC объединяет более 223 000 сотрудников в 157 странах. Более подробная информация представлена на сайте <http://www.pwc.ru/ru/about.html>