

2. Обзоры

2.1. OPENTALKS.AI-2022: ОБЗОР ПОСЛЕДНИХ ДОСТИЖЕНИЙ В ОБЛАСТИ АНАЛИЗА ЕСТЕСТВЕННОГО ЯЗЫКА И СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

М. А. Милкова, ЦЭМИ РАН, Москва

Журнал Цифровая экономика вновь публикует обзор некоторых докладов конференции OpenTalks.AI-2022 – независимой открытой конференции по искусственному интеллекту в России, объединяющей ведущих разработчиков, ученых и представителей бизнеса. На этот раз наша статья включает обзор ключевых достижений в области анализа естественного языка, историю развития сильного искусственного интеллекта, а также обсуждение места России в мире в области ИИ.

Главные достижения и тренды в NLP

Обзор главных результатов 2021 года в обработке естественного языка (NLP) представил Григорий Сапунов (Intento). Ключевое направление развития искусственного интеллекта базируется на использовании моделей-фундаментов (Foundation models), которые были представлены в 2021 году Стэнфордским университетом (Stanford Institute for Human-Centered Artificial Intelligence). Модели-фундаменты представляют собой набор типовых архитектур нейронных сетей, обученных на огромных массивах данных и имеющих широкие возможности для последующей адаптации их к широкому кругу конкретных задач. Иначе говоря, происходит некоторая гомогенизация моделей – модели-фундаменты позволяют решать задачи, для которых изначально модели не создавались. Описание же задачи происходит через демонстрацию примеров. К такого рода моделям относятся GPT-подобные модели, различные модели BERT, ELMO и др. Основным недостатком такого подхода является уязвимость модели: если в исходной модели были уязвимости – они останутся и во всех остальных.

Языковые модели непрерывно растут, постоянно ощущается давление в этой области. В настоящее время уровень человека побит, и наметилась тенденция на появление новых задач, которые нельзя было представить раньше.

Отметим некоторые интересные области применения NLP. Это разработка моделей для помощи в написании программного кода, позволяющего генерировать код для рутинных операций (например, Codex от OpenAI, alfaCode от DeepMind), развитие возможностей по генерации картинок через текстовое описание (например, DALL-E, CLIP от OpenAI), определение близости картинки к тексту (алгоритм ClipDraw, на основе кривых Безье), создание сервисов для помощи писателям (например, NovelAI),

В 2021 году продолжается развитие машинного перевода – в настоящее время машинным переводом покрывается более 100 тысяч языковых пар. Множество усилий ушедшего года также было направлено на развитие так называемых безтекстовых NLP, позволяющих работать с редкими, исчезающими языками, у которых либо нет письменности, либо имеется лишь небольшой по объему доступный корпус текстов, но сохранены аудиозаписи. В таком случае из аудиозаписи данные, минуя текст, попадают сразу в модель. Применение NLP здесь служит целям цифровизации языка, его сохранению.

Отдельным направлением развития является применение нейронных сетей в области биологии для описания белков, предсказания структуры белков.

В обзоре трендов NLP (Татьяна Шаврина, SberDevices, AIRI) отмечено, что NLP является своеобразным донором моделей для задач других доменов (компьютерное зрение, генерация кода, медицина). Также намечен тренд на мультимодальность (т.е. совместное использование в моделях текстов, изображений и др.).

В настоящее время существует несколько десятков больших моделей для русского языка и CV, которые в том числе доступны в открытых репозиториях: ruGPT-3 (продолжение текста, генерация контента), ruTS (энкодер-декодер русского языка), ruRoBERTa, ruBERT (энкодер русского языка); ruCLIP (сопоставление картинки и текста), ruDALL-Y (генерация изображений), RuDOLPH (видоизменение изображения по текстовому запросу, автоматическое описание картинки) и др.

Суперкомпьютеры

Отдельная сессия была посвящена суперкомпьютерам, в центре которой был доклад Дмитрия Монахова (Яндекс) о построении Яндексом собственных GPU-кластеров для распределенного машинного обучения. Построенные кластера машинного обучения позволяют осуществлять эффективное

энергоёмкое обучение больших моделей (Large models). Кластера заняли 19 (Chervonenkis), 36 (Galushkin) и 40 (Луарупов) места в рейтинге суперкомпьютеров Top500¹ в ноябре 2021 года.

Андрей Петров, Алексей Климов (SberCloud) также рассказали об опыте построения суперкомпьютеров в Сбербанке (Christofari, Christofari Neo – 72 и 43 места в рейтинге Top500 в ноябре 2021). Кроме того, на базе суперкомпьютера построена платформа ML Space для совместной работы команд над созданием и развертыванием моделей машинного обучения. Среди предпосылок к появлению суперкомпьютера и построению платформы на его основе отмечаются необходимость в эффективном использовании ресурсов, упрощение работы специалистов по анализу данных — для чего на платформе создаются преднастроенные шаблоны для распределенного обучения. Суперкомпьютер используется не только внутри Сбербанка, но может также быть доступен любому юридическому лицу после заключения соответствующего договора.

В целом, распределенное обучение может быть представлено тремя видами. Самое простое представляет из себя *параллелизм по данным* – данные разбиваются на фрагменты (батчи/эпохи), каждый фрагмент обучается на своем вычислителе, затем производится операция сбора. Данный вариант работает, когда сами веса нейронных сетей, оптимизатор влезает в память одного вычислителя. Однако бывают ситуации, когда модель настолько большая, что необходимо осуществлять *параллелизм по модели*, когда одна модель режется по слоям нейронной сети и вычисляется по отдельности. Наиболее сложным является вариант, когда обучают гигантские объемы данных (сотни гигабайт) – в данном случае сочетаются параллелизм по данным, по модели, а также pipeline-параллелизм (объединение нескольких операций обработки данных в единую модель) – так называемый *3d-параллелизм*.

Применение NLP в бизнесе

Отдельная секция была посвящена применению NLP в бизнесе. Свой доклад представил Илья Кузьминов (Институт статистических исследований и экономики знаний НИУ ВШЭ). Речь шла о разработанной в институте системе iFORA (Intelligent Foresight Analytics), предназначенной для анализа больших данных. Система агрегирует более 400 млн. документов разного типа: научные публикации, отчеты НИР, гранты, научные конференции, патенты, образовательные программы, вакансии, нормативные документы, профессиональные СМИ, научные проекты международных и национальных программ. iFORA позволяет оценивать степень представленности понятий и возникающих направлений, сравнивать представление по различным источникам, что дает понимание ландшафта технологий и зарождающихся рынков. Система позволяет проводить анализ рынка труда в связи с образовательными программами, выявлять наиболее значимые направления ИИ, востребованные навыки на рынке труда, строить матрицы применения технологий в различных секторах экономики, улавливать редкие и высокоинформативные понятия. Система работает с русским, английским и китайским языками.

Отдельный блок был посвящен применению NLP в медицине. Так, доклад Александра Гусева (Webiomed.AI) освещал вопросы анализа электронных медицинских карт для прогнозирования, а также общую динамику в области цифрового здравоохранения (см. Рис. 1). По опыту мировых компаний, применение NLP в этой области способно снижать смертность у пациентов высокого риска (на 20%, по данным Unity Health Toronto), сокращать число госпитализаций (на 10%, по данным Signify Health), экономить затраты медицинского центра (по данным Olive). Цель применения NLP здесь – в автоматическом извлечении из карт специальных признаков, которые направляются в прогнозные модели и позволяют выводить подсказки врачу (например, о возможном наличии у пациента заболевания, которое врачом не диагностировано). Ключевой проблемой здесь является нахождение ошибочных данных в картах (карты с ошибочными данными встречаются в 50% случаев), формирование качественных наборов данных реальной клинической практики.

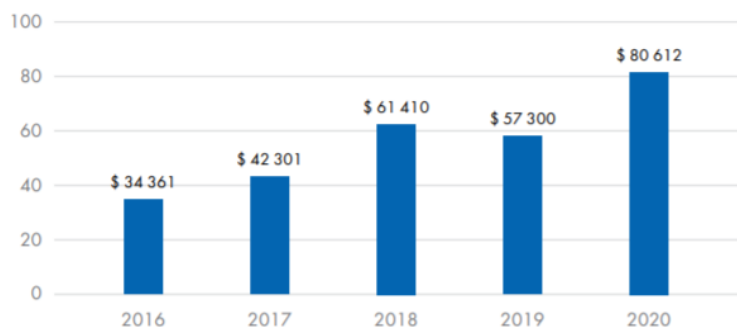


Рис. 1. Динамика инвестиций в продукты цифрового здравоохранения 2016-2020 гг. В 2020 г. был установлен абсолютный максимум инвестиций в цифровое здравоохранение, достигнув показателя 80.6 млрд. долл. США²

¹ Top500 —рейтинг 500 самых мощных вычислительных систем мира. Проект был запущен в 1993 году и публикует два раза в год актуальный перечень суперкомпьютеров. URL: <https://www.top500.org/lists/top500/list/2021/11/>

² Пугачев и др. (2021), со ссылкой на отчет CB Insights

Лоран Акопян и коллеги (iPavlov) представили систему скрининга деперсонализированных электронных медицинских карт с целью выявления редких генетических заболеваний (орфанные заболевания у детей до 7 лет). Выявление с помощью ИИ призвано сократить среднее время, необходимое для постановки диагноза с 7 до 2 лет и снизить детскую смертность на 50%. Основная задача здесь – это поиск сущностей биомаркеров с целью формирования векторного представления медицинской карты.

Ирина Жилина (Леруа Мерлен), совместно с Павлом Егоровым (GlowByte) представили проект по применению NLP в HR-аналитике в Леруа Мерлен. Прирост показателя удовлетворенности сотрудников ведет к положительному изменению индекса потребительской лояльности клиентов, таким образом HR-подразделение может влиять на коммерческий результат, улучшая опыт сотрудников. Обработка открытых вопросов анкет является трудозатратным процессом (занимает около 200 часов работы – 1.5 месяца работы сотрудника). Снизить затраты на обработку открытых вопросов (проведение мультимодальной кластеризации и анализ тональности) позволило применение тематического моделирования (модель Gibbs Sampling Dirichlet Multinomial Mixture, GSDMM), моделей Fast text, ruBert.

Сильный искусственный интеллект: от слепых мудрецов к общей теории интеллекта

На пленарной сессии Игорем Пивоваровым был представлен обзор текущего состояния области сильного искусственного интеллекта (Artificial General Intelligence, AGI). В данной статье мы передаем доклад максимально подробно, дополняя для удобства читателя необходимыми ссылками.

Диалог о сильном искусственном интеллекте стоит начать с того, что мы вообще понимаем под интеллектом. Прежде определим процессы обучения и понимания. Обучение – это способ накапливать информацию о результатах своего поведения и адаптировать поведение с какой-либо целью. Понимание – корректировка своих представлений, целей, картины мира и, в конечном счете, поведения в соответствии с новой информацией/событиями. Интеллект же стоит над обучением и представляет собой способность к долгосрочному планированию действий – умение ставить цели и достигать их (в условиях неопределенности с ограниченными ресурсами). Однако необходимо говорить также и о других составляющих агента, а именно, о психике, которая представляет собой субъективный мир агента, воспринимающего внешний мир, включающий комплекс субъективно переживаемых процессов и состояний (ощущения, память, мышление, мотивация, активность и др.), определяющих его поведение. Психика агента мотивирует его к обучению. Само по себе обучение будет бесцельно, а интеллект – это тот инструмент, которым агент будет пользоваться. Сознание – еще одна отдельная сущность, которая сегодня останется за рамками обсуждения.

Исходя из вышеизложенного сильный искусственный интеллект – это модель, вырабатывающая целенаправленное поведение для достижения любых целей, демонстрирующая понимание входной информации. Таким образом, здесь больше составляющих, чем просто «интеллект».

Тестирования для моделей AGI (например, один из первых тестов – тест Тьюринга) – направляют всю разработку. Никто не может определить, что такое интеллект однозначно. Часто это похоже на притчу о слепцах, которые ощупывают слона с разных сторон и выносят каждый свой вердикт. Сегодняшние тесты для AGI подобны ощупыванию слона с разных сторон. Люди, которые создают модели, удовлетворяющие этим тестам, как бы воссоздают некоторые части слона. Но как сделать тесты, которые могли бы воспроизвести всего слона?

Сначала рассмотрим существующие подходы к тестированию. В настоящее время можно выделить 3 блока тестов: языковые тесты (к которым относится тест Тьюринга, бенчмарк SuperGLUE и др.), тесты поведения (например, игры в Atari) и некоторые теоретические тесты, пока никак не реализуемые (например, Coffee test, который определяет, что робот обладает сильным искусственным интеллектом, если он сможет зайти в любой дом, зайти на кухню и сварить чашку кофе).

Изначальная позиция ученых состояла в том, что именно понимание языка – это путь к настоящему искусственному интеллекту. Долгое время мейнстримом был тест Тьюринга (Alan Turing, 1950). Однако сейчас мы понимаем, что, прохождение машины теста (машина может разговаривать неотличимо от человека) совсем не означает, что машина обладает интеллектом. Более современными являются другие тесты, например, тест на поведение PHIRE.AI, созданный Facebook, в котором агенту необходимо создать шар нужного размера в нужном месте так, чтобы дальше он с учетом гравитации падал, и в результате соприкоснулись два других шарика – синий и зеленый. Модель научилась ставить шарик, но понимания физики у нее нет. Существуют и другие тесты, однако все они лишь симулируют наличие интеллекта: «если нечто выглядит как утка, плавает как утка и крикает как утка, то, вероятно, это и есть утка».

Рассмотрим далее **подходы к моделированию интеллекта**. К ним относятся *когнитивные архитектуры, языковые модели, мультимодальные модели, модели поведения* и в итоге – *общая теория интеллекта*.

Немного отступая, сделаем ремарку о нейросетях. Искусственная нейронная сеть – гигантская система, которая прекрасно занимается аппроксимацией, внутри которой, очевидно, ни о каком интеллекте речи не идет. Лучшие нейросети сейчас – это те, что сконструированы по образу и подобию структур зрительной коры (Convolutional neural network, CNN). Они созданы, чтобы распознавать и классифицировать объекты, но не для того, чтобы моделировать целесообразное поведение.

1. *Когнитивные архитектуры.* Приверженцы такого подхода полагают, что главное в нашем мышлении и интеллекте – это способность рационально рассуждать, и именно это и надо моделировать, это – основа интеллекта. Когнитивные архитектуры представляют собой подход «сверху вниз», моделирующий основные когнитивные функции человека (например, память). Подход базируется на построении онтологии – набора объектов и связей между ними (например, в виде графа). Предполагается, что посредством построения онтологий возможно описывать картину мира. Пример такой архитектуры – система неаксиоматического вывода (Non-Axiomatic Reasoning System, NARS), которую представил и развивает Пэй Ван (Wang, 1995; Wang et al., 2018).

Другой интересный проект в этой области – проект Singularity.Net (Ben Goertzel) – подразумевает, что мы можем создать сильный ИИ как сумму множества слабых, работающих как сервисы в едином облаке. Иначе говоря, это некая метаонтология, где ее объекты – это слабые ИИ, а в центре находится также когнитивная архитектура – OpenCog (с онтологическим движком AtomSpace³) (Hart and Goertzel, 2008).

Когнитивные онтологии можно сравнить с картами, которые подобны миру, но не равны ему. Мы понимаем, все, что находится внутри, но всегда может найтись новый объект, который уже не попадет в текущую классификацию, что будет непрерывно требовать перестройки классификатора. Данное направление много лет было мейнстримом, но в последнее время сильно угасло и, видимо, за ним не будет будущего.

2. Сейчас мейнстримом являются гигантские глубокие нейросети и в первую очередь – *языковые модели.* Языковые модели – это как утверждение другого слепого мудреца, который говорит, что главное в нашем мышлении и интеллекте – это язык. Он не только средство коммуникации, на языке мы «мыслим». Именно язык является фундаментом и именно его надо моделировать. Моделирование языка начиналось с модели word2vec (Mikolov, et al., 2013), позволяющей переводить слова в векторную форму, оценивать близость слов и таким образом имитировать понимание как близость слов друг к другу.

Сейчас же моделирование происходит с помощью архитектур Transformers (разработанные в компании OpenAI), которые на основе механизма Attention считывают слова совместно с их контекстом и, пройдя множество итераций, определяют, что может значить слово в контексте остальных (Vaswani et al., 2017). Аналогично устроены и визуальные трансформеры, которые режут картинку на части и позволяют объяснить, что на ней, используя контекст.

Пиком развития глубоких нейросетей сегодня является модель GPT-3 – она не просто обучается на тексте, но и может генерировать его. Сеть GPT-3 способна отвечать на вопросы, и при этом ее способности выше, чем у некоторых людей⁴.

Мнения о GPT-3 разнятся от «вполне возможно, что сегодняшние нейронные сети немного обладают сознанием»⁵ до таких, которые смотрят на GPT-3 как на некоторого стохастического попугая, который обучен на таком гигантском объеме текстов, который человеку не по силам прочитать (Bender et al., 2021).

Однако, нам кажется, что глубокие сети, конечно, не обладают никаким сознанием. Они предоставляют лишь связки между словами, имеют некую ассоциативную память, а отнюдь не картину мира. Формирование картины мира нужно агенту для каких-то целей, а обучение GPT-3 бесцельно. От ответов GPT-3 для нее ничего не зависит. Возможно, если бы от ее ответов зависела ее же жизнь (если бы она могла быть), может, она бы отвечала совсем по-другому.

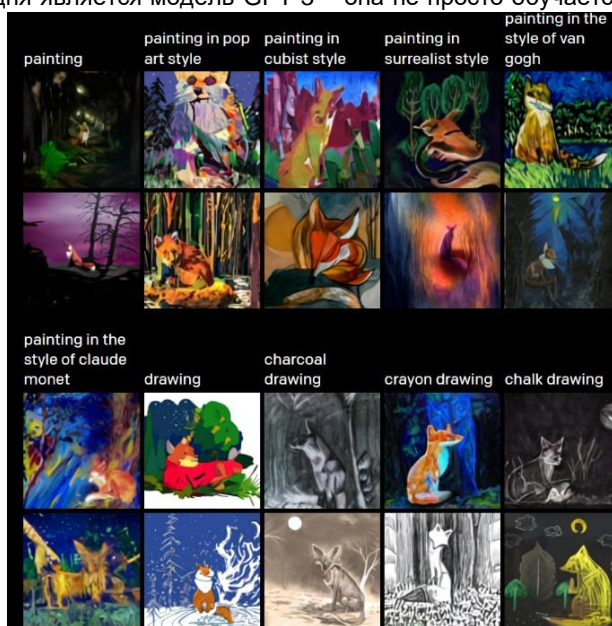


Рис. 2. Фрагмент результатов работы DALL-E по запросу «a painting of a fox sitting in a forest at night»⁶.

3. После языковых моделей возникло еще одно движение – это *мультимодальные модели.* Главное в нашем мозге – это ассоциативность, связь между образами, ощущениями, словами языка. Поэтому одного языка недостаточно. Нужно моделировать именно ассоциативность, вместе подавать

³ Подробнее об OpenCog AtomSpace, см. <https://wiki.opencog.org/w/AtomSpace>

⁴ Разговор биолога Александра Панчина и GPT-3: «Мои враги – природа, энтропия и смерть»: искусственный интеллект о Боге, гомеопатии и старении.» URL: <https://scinquisitor.livejournal.com/183774.html>

⁵ Публикация в Твиттере Илу Sutskever, главный научный сотрудник OpenAI, 10.02.2022.

⁶ Источник: <https://openai.com/blog/dall-e/>

модели и текст, и картинки, видео. Примером мультимодальной является модель DALL-E, которая берет модель GPT-3 и дообучает ее на картинках (Ramesh et al., 2021).

Действительно, некоторые функции нашей памяти проявляются в этой архитектуре. Однако, мультимодальность – это ассоциативная память, а для создания картины мира должна быть цель.

4. Согласно же другому подходу, главное в нашем мышлении и интеллекте — это то, что они управляет *поведением*. В том числе осуществляя планирование и выполнение действий. Обучение в данной области опирается на концепцию *обучения с подкреплением* (Reinforcement learning, RL). Агент (модель) учится на неразмеченных данных (например, в виртуальной среде), получая: вознаграждение (positive reward) за правильные действия, наказание (negative reward) за неправильные действия.

Примерами такого рода моделей являются DQN – модели, которые выигрывают игры Atari у человека (Mnih et al., 2013); модель Open AI Five – научилась играть в стратегическую игру Dota2, причем команда из 5 взаимодействующих между собой ИИ выиграла у команды из 5 человек.

Еще один пример – игра Hide and Seek (Baker et al., 2020), в которой одна группа агентов ищет другую группу агентов, которые от них прячутся, используя разнообразные объекты – кубики, укрытия. Через много миллионов итераций одна группа научилась загромождать кубиками, другая группа, в свою очередь, научилась ставить доску и забираться по ней, преодолевая укрытия.

RL модели помогают моделировать поведение, что является более перспективной идеей, чем предыдущие. Однако главный вопрос остается тем же – есть ли здесь интеллект? Или RL агент просто учится рефлексам? Вопрос остается открытым.

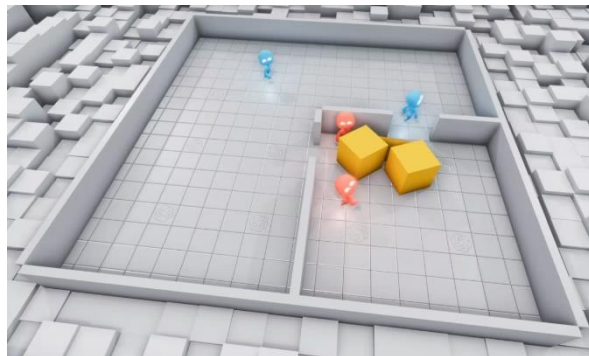


Рис.3. Иллюстрация игры Multi-Agent Hide and Seek⁷

Общая теория интеллекта. Игорь Пивоваров, будучи физиком-теоретиком, подчеркивает, что не нужно моделировать отдельные функции и потом складывать их вместе. Нужно создать общую теорию интеллекта, из которой дальше вытекут все детали архитектуры, «сверху-вниз». И в последнее время появились зачатки такой теории.

Карл Фристон сформулировал принцип свободной энергии (Free Energy Principle) (Allen and Friston, 2018). Есть фундаментальный физический принцип – принцип наименьшего действия – и его можно использовать применительно к живым организмам и интеллекту.

Принцип наименьшего действия. Его первый практический вид сформулировал Пьер Ферма, и он относился к преломлению луча света при переходе из одной среды в другую: луч преломляется под таким углом, чтобы суммарное время в пути фотонов было минимальным. Аристотель в свое время писал, что в природе ничто не делается просто так – все делается с минимальными усилиями. Эйнштейн говорил, что всю общую теорию относительности можно было вывести из принципа наименьшего действия. Таким образом, принцип наименьшего действия может быть приложен к любому физическому закону, в терминах минимизации/максимизации какой-то величины. Система всегда выбирает экстремальный путь для достижения результата. Но почему мир так устроен? Можно ли связать жизнь и энтропию (как меру беспорядка)? В мире энтропия не убывает (согласно второму закону термодинамики), или, иначе – хаос нарастает. А жизнь – это нечто, что противостоит этому беспорядку. Можно сказать, что жизнь – это стационарное поддержание устойчивого неравновесия (гомеостаз). Жизнь сопротивляется энтропии. Можно ли найти принцип наименьшего действия, который бы это объяснял – что минимизирует жизнь?

Именно в этом русле начал работу Карл Фристон. Задача организма – это предсказание неожиданности. Наш интеллект помогает нам прогнозировать так, чтобы минимизировать неожиданности. Задача мозга – избегать неожиданности (неопределенности, риски). Минимизация неожиданности – минимизация «свободной энергии» – теоретической информационной величины, характеризующей ошибку предсказания.

Фристон сформулировал понятие *предиктивного кодирования*, согласно которому мозг создает иерархическую картину мира на базе сенсорно-моторных сигналов, на основании этой картины мира мозг старается предсказать следующее состояние (и делает это на каждом уровне иерархии). На каждом уровне есть ошибка предсказания, и каждый следующий уровень стремится минимизировать ошибку предыдущего уровня. Задача мозга – минимизировать ошибку предсказания, а предсказываем мы то, где нам надо находиться, чтобы сохранить гомеостаз, чтобы выжить (Ondobaka et al., 2017). Существуют два ключевых исследователя, занимающихся этой темой – это Джефф Хокинс (Hawkins, 2021) и Сергей Шумский (Шумский, 2020).

⁷ Источник: <https://openai.com/blog/emergent-tool-use/>

Таким образом, последним направлением развития сильного искусственного интеллекта является моделирование психики. Именно психика мотивирует человека к обучению, и моделирование интеллекта, который находится над этими процессами, без моделирования психики невозможно. «Интеллект без психики – как инструмент без мастера», – заключает Игорь Пивоваров⁸.

Этика ИИ

Отдельный доклад об этике разработки и применения ИИ представил Александр Крайнов (Яндекс). Отличие в оценке действий ИИ и человека заключается в том, что решение одного человека в повседневной жизни, как правило, не очень сильно влияет на других. Технологии ИИ же масштабируют решение одного человека (команды) на всех. Речь идет, к примеру, о работе социального рейтинга, управлении беспилотным автомобилем и т.п. Кроме того, в этике человека никогда не ставился вопрос о том, что число происшествий делилось на число происшествий. 2021 год ознаменовался созданием национального кодекса этики в сфере ИИ. Кодекс разработан Альянсом в сфере искусственного интеллекта⁹ и рассматривает вопросы о прозрачности методов, справедливости решений, формировании среды доверия и др.

Место России в мире в искусственном интеллекте

Country	Implementation			Innovation			Investment		Total rank
	talent	infrastructure	operating environment	research	development	government strategy	commercial		
United States of America	1	4	35	1	1	17	1	1	1
China	24	1	6	2	2	2	2	2	2
United Kingdom	3	23	24	5	11	11	4	3	3
Canada	7	15	5	10	10	1	6	4	4
Israel	5	29	14	7	9	43	3	5	5
Singapore	4	8	36	4	14	15	5	6	6
South Korea	28	6	32	12	3	7	15	7	7
The Netherlands	6	9	10	15	8	33	18	8	8
Germany	11	13	30	6	12	10	8	9	9
France	9	14	17	16	16	5	10	10	10
Australia	15	38	41	8	4	12	13	11	11
Ireland	8	5	31	18	7	25	25	12	12
Finland	16	22	18	17	20	8	21	13	13
Denmark	13	18	11	13	16	18	26	14	14
Luxembourg	18	3	33	25	17	30	20	15	15
Japan	26	7	48	19	5	21	12	16	16
India	2	69	19	25	6	38	11	17	17
Switzerland	14	11	54	3	13	58	9	18	18
Sweden	10	17	34	11	22	47	22	19	19
Hong Kong	22	2	44	9	23	50	17	20	20
Spain	21	19	23	26	29	4	28	21	21
Austria	23	34	21	18	21	19	29	22	22
Estonia	19	37	9	14	30	20	7	23	23
Taiwan	35	12	50	14	16	58	32	24	24
Norway	12	16	58	21	28	37	24	25	25
Saudi Arabia	55	25	1	21	25	3	19	26	26
Belgium	27	32	37	20	18	32	16	27	27
Poland	30	24	2	16	31	16	34	28	28
Slovenia	33	21	4	24	18	14	46	29	29
New Zealand	17	26	7	32	30	43	33	30	30
Italy	37	33	13	22	24	34	31	31	31
Russia	34	41	51	30	10	5	38	32	32

Рис. 4. Рейтинг The Global AI Index 2021¹⁰

В первом случае ключевым механизмом является монополия на инновацию, а получение экономической отдачи происходит с начальной стадии внедрения. Во втором случае ключевой механизм – это

Завершилась конференция открытым разговором на тему «Место России в мире в искусственном интеллекте», что представляет для нас сейчас особый интерес. В дискуссии участвовали руководитель OpenTalks.AI, Игорь Пивоваров, Александр Крайнов (Яндекс), Олег Сиротюк (Huawei), Дмитрий Конягин (NVIDIA), Сергей Карелов (IT-эксперт, ведущий канала «Малоизвестное интересное»), Константин Воронцов (МФТИ).

Дискуссию начал Сергей Карелов с тезисом «Место России в ландшафте искусственного интеллекта: нас там нет». Согласно композитным индексам, Россия является аутсайдером (см., например, Stanford HAI Index 2021 – Россия занимает последнее, 26 место; The Global AI Index 2021 – Россия на 32 месте).

Существуют и другого типа индексы – например, AI2000 – где область ИИ поделена на 20 областей, в каждой из которых выделено по 100 самых продвинутых ученых, оценивающих достижения стран. Россия в данном рейтинге не представлена.

Однако, как отметил Александр Крайнов, составление рейтингов зачастую учитывает такие параметры, по которым Россия априори не может быть хорошо представлена. Так, например, в некоторых рейтингах учитывается число пользователей в LinkedIn, указавших сферу своей специализации как ИИ или же учет количества специалистов по ИИ на душу населения. Также в рейтингах не учитывается такой показатель, как массовое внедрение собственных технологий. И в этом разрезе Россия – одна из немногих стран после США, у которых есть свой поисковик, социальная сеть, беспилотный автомобиль.

Сергей Карелов отметил, что существует два класса технологий: технологические продукты (ТП) (автомобиль, самолет, смартфон) и технологии общего назначения (ТОН) (электричество, компьютер, ИИ). У этих двух классов принципиально разные механизмы и технология развития.

⁸ Подробнее об устройстве интеллекта и сознания, см. И. Пивоваров (2021).

⁹ Альянс в сфере искусственного интеллекта. URL: <https://a-ai.ru>

¹⁰ Источник: <https://www.tortoisemedia.com/intelligence/global-ai/>

широта проникновения, а получение экономической отдачи отсрочено на последнюю стадию внедрения. Для того, чтобы быть лидером по технологическим продуктам, принципиально важно быть мировым лидером по R&D, по числу статей и патентов, по числу топовых ученых и инженеров, создающих ТП. Именно эти параметры и меряют композитные индексы. Для ТОН все эти индексы не годятся. Для ТОН важно совершенствование инфраструктуры и оргподдержки, стимулирующих создание инновационных цепочек взаимодополняющих применений ТОН в большинстве секторов экономики. Важно обеспечение большого пула инженерных специалистов среднего и нижнего уровня, ориентация на расширение их возможностей.

США изначально были и есть лучше других стран подготовлены к прогрессу ИИ как ТОН. Руководство Китая в явном виде ориентирует национальную стратегию как на стратегию ТОН. В России это совсем не так, однако она могла бы участвовать в этой гонке. Но для этого необходимо менять стратегию, организовывать что-то типа «интегрированной стратегии ИИ-ГОЭЛРО (программа всеобщей электрификации) плюс ИИ-Ликбез (увеличение специалистов среднего и нижнего уровня)». Широта применения зависит от инфраструктуры. И в этом плане место России в ИИ как ТОН может быть очень неплохим, но для этого необходимо менять стратегию.

Константин Воронцов отметил необходимость создания в России среды, где могли бы рождаться новые идеи. Такую среду, комфортную для мозгов, необходимо конструировать. Должны рождаться коллективы, способные делать топовые статьи, а для этого они должны понимать новые задачи. Одним из вариантов поднятия интереса к науке ИИ может быть создание некоего сервиса, наподобие Kaggle¹¹, в котором задачи различных прикладных областей общедоступны и популярно рассказаны, а анализ данных является точкой входа в эти области.

Дмитрий Конягин отметил важность создания доступной вычислительной инфраструктуры – если нет вычислительных мощностей, то работа больших моделей невозможна. У Яндекса, Сбера есть свои мощности, однако необходимо создавать центры коллективного пользования.

Перспективы российского железа также не выглядят радужными. В России, впрочем, как и в Европе, нет изготовителей чипов, компании занимаются лишь их дизайном (дизайн-бюро)¹². Важно и то, что все ставит во главу угла – железо, но сделать из железки продукт – это для российских компаний представляет собой «ту самую часть айсберга, которая находится под водой» (Олег Сиротюк).

Таким образом, особую важность представляет вопрос технологического суверенитета страны – в случае объявления эмбарго России по геополитическим причинам, плана Б у России в части искусственного интеллекта, увы, нет. Константин Воронцов подытожил, что развитие ИИ является социально-политическим вопросом и единственным позитивным взглядом на эту ситуацию может быть только взгляд на эволюцию технологий с точки зрения биологической эволюции, где рациональный путь развития – это обособление, которое не исключает сотрудничество, но подразумевает разработку своей модели развития.

Список литературы

14. Пивоваров, И. (2021). ТЕРБИДЕНЬ – простая модель сложного мира. Ridero. С. 130 с.
15. Шумский, С.А. (2020). Машинный интеллект. Очерки по теории машинного обучения и искусственного интеллекта. Издательство РИОР. С. 340 с.
16. Allen, M., Friston, K.J. (2018). From cognitivism to autopoiesis: towards a computational framework for the embodied mind // *Synthese*, 195 (6): 2459-2482
17. Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., Mordatch, I. (2020). Emergent Tool Use From Multi-Agent Autocurricula // arXiv:1909.07528
18. Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? // *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March 2021, pp. 610–623 <https://doi.org/10.1145/3442188.3445922>
19. Hart, D., Goertzel, B. (2008). OpenCog: A Software Framework for Integrative Artificial General Intelligence // *Proceedings of the First AGI Conference, AGI 2008*, pp. 468-472.
20. Hawkins, J. (2021). *A Thousand Brains: A New Theory of Intelligence*. Basic Books: New York. 288 P.
21. Mikolov, T., Kai C., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space // arXiv:1301.3781
22. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning // arXiv:1312.5602
23. Ondobaka, S., Kilner, J., Friston, K. (2017). The role of interoceptive inference in theory of mind // *Brain and Cognition*, 112, pp. 64-68
24. Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., Sutskever, I. (2021). *Ze-ro-Shot Text-to-Image Generation* // arXiv:2102.12092
25. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin,

¹¹ Kaggle — система организации конкурсов по исследованию данных, а также социальная сеть специалистов по обработке данных и машинному обучению. Источник: Википедия.

¹² На момент написания данной статьи стало известно, что 6 марта 2022 компания Nvidia (разработчик и основной поставщик чипов) приостанавливает поставки в Россию.

- I. (2017). Attention Is All You Need. arXiv:1706.03762v5
26. Wang, P. (1995). Non-Axiomatic Reasoning System: Exploring the Essence of Intelligence // PhD Thesis, Indiana University. 181 P.
27. Wang, P., Li, X., Hammer, P. (2018). Self in NARS, an AGI System // Frontiers in Robotics and AI, 5. <https://doi.org/10.3389/frobt.2018.00020>.

References in Cyrillics

9. Pivovarov, I. (2021). TERBIDEN is a simple model of a complex world. Ridero. p. 130 p.
10. Shumsky, S.A. (2020). Machine intelligence. Essays on the theory of machine learning and artificial intelligence. RIOR Publishing House. p. 340 p.

Милкова Мария Александровна – научный сотрудник лаборатории экспериментальной экономики ЦЭМИ РАН

Ключевые слова

Искусственный интеллект, сильный искусственный интеллект, моделирование интеллекта, NLP, анализ естественного языка

Maria Milkova. OpenTalks.AI-2022: An overview of recent advances in natural language processing and artificial general intelligence

Keywords

Artificial intelligence, artificial general intelligence, intelligence modeling, NLP, natural language processing

DOI: DOI: 10.34706/DE-2022-01-08

JEL classification: D83 – Поиск • Обучение • Информация и знания • Взаимодействие • Мнение • Неосведомленность

Abstract

Digital Economy journal republishes a review of some reports of the OpenTalks.AI-2022 conference, an independent open conference on artificial intelligence in Russia, bringing together leading developers, scientists and business representatives. This time, our article includes an overview of key advances in natural language processing, the history of the development of artificial general intelligence, and a discussion of Russia's place in the world in the field of AI.