



АВТОРСКАЯ АНАЛИТИЧЕСКАЯ СТАТЬЯ

Экономика заимствованного интеллекта

Как упаковка чужого разума превратилась в индустрию, а подлинное изобретение интеллекта стало чужаком среди тех, кто присваивает себе это звание

В 1984 году небольшая нидерландская компания начала работу в сарае с протекающей крышей неподалёку от производственных площадок Philips в Эйндховене (Eindhoven). Она собиралась разрабатывать установки фотолитографии (Photolithography) для производства полупроводников (Semiconductors). 1

Её звали **ASML**.

Четыре десятилетия спустя ASML стала единственной компанией, поставляющей коммерческие системы литографии в экстремальном ультрафиолете (Extreme Ultraviolet Lithography — EUV) — одни из самых сложных машин в истории современной промышленности.

Внутри такой установки вылетают микроскопические капли расплавленного олова. Их поражают последовательные лазерные импульсы — пятьдесят тысяч раз в секунду. Вещество превращается в плазму (Plasma), излучающую свет с длиной волны 13,5 нанометра. Затем свет проходит через сверхточную оптическую систему и участвует в нанесении микроскопических схем на кремниевые пластины. 2

Эта машина — победа не одной лишь механики. Она возникла на пересечении оптики (Optics), физики плазмы (Plasma Physics), материаловедения (Materials Science), теории и техники управления (Control Engineering), программного обеспечения и прецизионной метрологии (Precision Metrology).

Ещё интереснее то, что ASML изобрела не все эти науки и не все компоненты своей установки. Компания выстроила специализированную сеть поставщиков и партнёров, среди которых немецкая ZEISS, отвечающая за оптику. Американская Cymer сотрудничала с ASML в разработке источника EUV, а в 2013 году ASML приобрела её, чтобы ускорить эту разработку. 3 4

Но именно ASML обладала инженерным знанием, необходимым, чтобы собрать эти компоненты в систему, которую прежде никто в мире не умел производить в таком виде.

Нидерланды стали силой в этой отрасли не потому, что купили новейшие установки для выпуска чипов, а потому, что выросшая на их земле компания начала производить машину, без которой не могут обойтись изготовители самых передовых микросхем.

В этом и состоит разница между обладанием технологией как товаром и обладанием знанием, благодаря которому её вообще можно произвести.

Пока ASML десятилетиями накапливала знания и тем самым выстраивала своё место в мировой экономике, в эпоху искусственного интеллекта возник другой тип компаний. Они стремятся занять столь же значимое положение, но выбирают путь куда короче: берут интеллект в аренду у другой компании, накрывают его изящным интерфейсом, а затем пишут на своём сайте, что разработали собственную технологию искусственного интеллекта (Proprietary AI Technology).

Возможно, впервые для перехода из пользователей технологии в её изобретатели в некоторых кругах оказалось достаточно поменять название продукта на экране.

Фабрики упаковки в эпоху арендованных умов

Генеративный искусственный интеллект (Generative AI) открыл историческую возможность для создания приложений. Одновременно он породил не менее процветающий рынок по производству технологических титулов.

Сегодня стартап может взять большую языковую модель (Large Language Model — LLM), разработанную OpenAI или Anthropic, подключиться к ней через программный интерфейс (Application Programming Interface — API), добавить генерацию с

дополненной выборкой (Retrieval-Augmented Generation — RAG), несколько инструкций (Prompts) и средства вызова инструментов (Tool Calling).

Немного оркестрации рабочих процессов (Workflow Orchestration) — и продукт уже выглядит как интеллектуальный агент (AI Agent): читает документы, подводит итоги совещаний, отвечает клиентам, составляет отчёты и взаимодействует с некоторыми корпоративными системами.

До этого момента перед нами вполне понятная разработка программного продукта. Она может быть полезной и прибыльной.

А затем происходит занятное превращение.

Компания начинает говорить о «собственном движке искусственного интеллекта», «передовой архитектуре рассуждений» и «проприетарных интеллектуальных моделях». Порой в своём самоописании она возвышается до компании глубоких технологий (Deer Tech), хотя способность, о которой идёт речь, по существу находится внутри GPT, Claude или другой коммерческой модели.

Они изобрели не разум. Они изобрели для него новое имя — и потребовали, чтобы рынок признал их научное отцовство.

У некоторых подобных компаний весь «движок искусственного интеллекта» состоит из ключа доступа (API Key), комплекта разработчика (Software Development Kit — SDK) и инструкций, адресованных модели, которая размещена у другой стороны.

И всё же основатель с воодушевлением рассказывает нам о своей когнитивной архитектуре, способности к рассуждению и технологической независимости.

Здесь трудно удержаться от иронии. Когда поставщик модели меняет цены на токены (Token Pricing), меняется экономика мнимого изобретения. Когда прерывается связь с сервисом, исчезает большая часть его интеллекта. Когда поставщик выпускает новую функцию, месяцы «революционной разработки» превращаются в бесплатную возможность, доступную всем.

В этой индустрии компания может утром объявить о своей интеллектуальной независимости, а вечером получить ошибку: **429 Too Many Requests**.

Сколько же независимых умов перестают думать, когда заканчивается месячная квота токенов.

В технологической среде это явление называют оболочками над API моделей (API Wrappers), а если точнее — оболочками над GPT (GPT Wrappers).

Они могут принимать форму программного обеспечения как услуги (Software as a Service — SaaS), платформ агентов (Agent Platforms) или отраслевых ИИ-приложений (Vertical AI Applications) для медицины, права, образования и финансовых услуг.

Смена отрасли не меняет происхождения способности.

Медицинский продукт, использующий Claude для анализа документа, не становится оригинальной медицинской моделью лишь потому, что его интерфейс одет в белое. Юридический агент, вызывающий GPT, не превращается в открытие в области юридического рассуждения только потому, что научился писать записки официальным языком.

Даже тысяча документов в базе поиска или подключение агента к десяти корпоративным системам ещё не означают, что компания изобрела языковую модель или новый алгоритм обучения.

Возможно, она создала превосходный продукт. Но превосходный продукт и изобретение в искусственном интеллекте — разные вещи.

Проблема не в использовании чужих технологий. Вся история промышленности строится на накоплении изобретений и взаимном использовании достигнутого.

Проблема начинается тогда, когда зависимость от чужого знания превращают в заявление, будто это знание создал ты сам.

В традиционной экономике мы называем это импортом, интеграцией или разработкой продукта. В экономике искусственного интеллекта некоторые уже называют то же самое «передовыми исследованиями и разработками».

И тогда перед нами уже не просто маркетинговое преувеличение, а намеренное размывание смысла инновации и характера той отрасли, которую мы хотим построить.

Рынок «лимонов», торгующий гениальностью

В 1970 году американский экономист Джордж Акерлоф (George Akerlof) опубликовал знаменитую работу *The Market for Lemons*. Она стала одним из основополагающих текстов для понимания информационной асимметрии (Information Asymmetry) на рынках.

Акерлоф исследовал, почему рынок может оказаться неспособным отличить хороший товар от плохого, когда продавец знает то, чего не знает покупатель. Такая ситуация приводит к отрицательному отбору (Adverse Selection), и рынок всё хуже вознаграждает подлинное качество.

Мне кажется, перед нами современная версия той же проблемы. Только товаром теперь служит не подержанный автомобиль, а притязание на обладание интеллектом.

Перед инвестором стоят две команды.

Первая потратила год на изучение архитектур моделей (Model Architectures), совершенствование алгоритмов обучения (Training Algorithms), эксперименты со стратегиями управления памятью и разработку новых способов обработки языка.

Вторая потратила несколько недель на красивый интерфейс поверх коммерческой модели.

Обе команды показывают свои продукты. Ответы выглядят безупречно, демонстрации — похожими.

Более того, у второй команды показ может оказаться красивее, быстрее и убедительнее для комиссии.

Многие инвестиционные презентации проверяют видимый результат, не разбираясь в происхождении технической способности (Technical Provenance). Поэтому обе стороны получают одно и то же определение: две передовые компании в области искусственного интеллекта.

Здесь и совершается маленькое экономическое преступление, которое растёт с каждым повторением.

Когда рынок не умеет отличить тех, кто производит знание, от тех, кто его арендует, он начинает вознаграждать умение изображать инновацию вместо умения её создавать.

Естественно, компании, способные дёшево устраивать впечатляющие показы, множатся, а исследовательские лаборатории кажутся медленнее, рискованнее и не столь эффектными в краткосрочной перспективе.

Исследователь может месяцами проверять гипотезы, пытаясь снизить расход памяти на 20%. В то же время другой основатель за несколько дней покажет агента, говорящего на пяти языках, благодаря модели, к разработке которой он не имел отношения.

Первый, возможно, ближе к созданию технологии, которая нужна миру. Второй — к созданию продукта, который нужен клиенту.

У каждого своя экономика.

Но смешение этих двух типов деятельности посылает рынку опасный сигнал: **не трать силы на разработку разума; научись продавать его под своим именем.**

Такой сигнал искажает не только решения инвесторов. Он меняет устремления целого поколения исследователей и инженеров.

Рынки не просто вознаграждают таланты. Они ещё и учат людей, кем стоит стремиться стать.

Если путь к титулу «лидера в искусственном интеллекте» чаще проходит через упаковку GPT, чем через разработку алгоритма, архитектуры или метода обучения, нет ничего удивительного в том, что компании-интерфейсы множатся, а серьёзные лаборатории всё труднее найти.

Тогда наша беда не в нехватке компаний.

Наша беда — в избытке притязаний на оригинальность и дефиците самой оригинальности.

Почему создатель модели поглотит своих приёмных детей

В исследованиях экономики платформ (Platform Economics) существует стратегия, известная как охват платформой смежных рынков (Platform Envelopment). Её рассматривали Томас Айзенман (Thomas Eisenmann), Джеффри Паркер (Geoffrey Parker) и Маршалл Ван Альстайн (Marshall Van Alstyne). [6](#)

Это происходит, когда крупная платформа распространяется на соседние функции и рынки, пользуясь своей технической инфраструктурой, пользовательской базой и способностью объединять сервисы.

Сегодняшняя ситуация в искусственном интеллекте напоминает мне именно такую динамику.

В марте 2025 года OpenAI представила [Responses API](#) и комплект [Agents SDK](#) со встроенными средствами поиска в интернете и файлах и использования компьютера, а также SDK для координации агентов.

Прежде стартапы продавали эти функции как самостоятельные продукты или отличительные технические возможности. Теперь они постепенно становятся возможностями самой платформы, которой принадлежит модель.

Уравнение здесь беспощадно в своей простоте.

Если ваше конкурентное преимущество — функция, которую поставщик арендуемого вами интеллекта может добавить одним обновлением, вы строите не технологическую крепость. Вы временно арендуете участок на чужой земле.

Ваш продукт сегодня может быть успешен. Но успех ещё не означает, что вы контролируете условия его дальнейшего существования.

Чем совершеннее становятся фундаментальные модели (Foundation Models), чем лучше они умеют пользоваться инструментами (Tool Use), выполнять многошаговые рассуждения (Multi-step Reasoning) и действовать с большей самостоятельностью (Agentic Execution), тем меньше пространства компании-оболочки могут продавать как исключительную добавку.

Некоторые компании прикладного ПО останутся коммерчески успешными: они понимают клиентов, умеют интегрировать системы или выстроили сильные каналы распространения.

Но это достоинства программного продукта и бизнеса, а не свидетельство того, что компания изобрела интеллект, которым пользуется её продукт.

Компания, чьим главным техническим достижением было сокрытие имени исходной модели, однажды столкнётся с необходимостью скрывать куда более неловкую правду: модель больше не нуждается в ней самой.

Мне представляется, что такое размежевание станет одним из важнейших рыночных сдвигов ближайших лет.

Экономическая ценность многих универсальных оболочек над ИИ (Generic AI Wrappers) будет снижаться. Одновременно возрастёт ценность компаний, владеющих оригинальными алгоритмами, более эффективными технологиями, специализированными моделями либо накопленными отраслевыми данными и знаниями, которые невозможно арендовать одним программным вызовом.

В то же время компании прикладного уровня окажутся под всё большим давлением: им придётся доказывать, что их ценность выходит за пределы функций базовых платформ.

Тем, кто представляет каждого интеллектуального агента как научное открытие, ближайшие годы, вероятно, будут меньше прощать маркетинговое красноречие.

В технологии одно обновление способно разрушить инвестиционный тезис, который строили годами.

Настоящие изобретения происходят по ту сторону экрана

Чтобы понять разницу между созданием интеллекта и созданием интерфейсов к нему, не обязательно рассматривать самые крупные языковые модели в мире.

Иногда изобретение глубже и куда менее зрелищно.

В 2022 году на конференции по нейронным системам обработки информации (NeurIPS) исследователь Три Дао (Tri Dao) и его коллеги представили работу FlashAttention.

Их целью было не создание нового чат-бота, а решение проблемы реализации механизма внимания (Attention Mechanism), используемого в архитектуре трансформеров (Transformers).

Исследователи обнаружили, что значительная часть замедления связана не с самими вычислениями, а с перемещением данных между памятью с высокой пропускной способностью (High Bandwidth Memory — HBM) и быстрой статической оперативной памятью внутри процессора (Static Random-Access Memory — SRAM).

С помощью алгоритма, учитывающего стоимость ввода-вывода (IO-Aware Algorithm) и разбивающего операции на блоки (Tiling), они сократили перемещение данных и повысили производительность точного внимания (Exact Attention). В одном из экспериментов, описанных в работе, выполнение было до трёх раз быстрее.

Они не добавили новую кнопку.

Они изменили способ выполнения базовой математической операции, и этим знанием затем воспользовались самые разные системы в индустрии моделей.

В 2023 году исследователи Калифорнийского университета в Беркли (UC Berkeley) представили систему vLLM и алгоритм PagedAttention.

Источником вдохновения послужила старая идея из операционных систем (Operating Systems): виртуальная память и её постраничная организация (Virtual Memory Paging).

Вместо того чтобы впустую занимать большие области памяти кэшем ключей и значений внимания (Key-Value Cache — KV Cache), исследователи разработали более эффективный способ управления этими данными при выполнении логического вывода (Inference).

Опубликованные эксперименты показали рост пропускной способности (Throughput) в два-четыре раза по сравнению с конкретными базовыми системами при сопоставимой задержке (Latency).

Сравните эти работы с компанией, которая заявляет о создании платформы искусственного интеллекта, потому что подключила коммерческую модель к базе данных.

В обоих случаях есть программное обеспечение.

Но в первом случае появляется новое знание о том, как работают модели. Во втором используется уже существующее знание.

В этом и состоит сущность глубинных технологий (Deep Tech).

Изобретение не обязательно требует обучения огромной модели с нуля (Training from Scratch). Оно может заключаться в новой архитектуре (Novel Architecture), алгоритме обучения, механизме сжатия модели (Model Compression), дистилляции знаний (Knowledge Distillation), способе адаптации моделей или их продолженного обучения (Continual Pretraining), а также в открытии, которое снижает расход ресурсов без потери качества.

Наука движется вперёд не потому, что каждое утро заново изобретает мир. Движение возникает, когда обнаруживается новая связь, снимается старое ограничение или разрозненные знания соединяются в конструкцию, открывающую прежде недоступные возможности.

Поэтому я различаю две внешне похожие компании.

Одна спрашивает: как заставить существующую модель выполнить задачу нашего клиента?

Другая спрашивает: как развить технологию, чтобы эта задача стала выполнимой способом, которого прежде не существовало?

Первая ищет продукт.

Вторая, возможно, строит целую отрасль.

Где создаётся техническая ценность

Критерий	Приложение на базе модели	Вклад в саму технологию
Вопрос	Как с помощью модели решить задачу клиента?	Как улучшить технологию, благодаря которой решается задача?
Предмет работы	Интерфейс, интеграция систем и рабочие процессы	Архитектура, алгоритмы, обучение и оценка
Возможная ценность	Полезный продукт и отраслевое или коммерческое преимущество	Знания и технические активы, на которых могут строиться другие продукты

Misraj и Kawn Lab: от потребления интеллекта к производству знания

Здесь я подхожу к тому, что считаю самым важным для будущего искусственного интеллекта в Саудовской Аравии.

Нам нужны компании, которые видят в искусственном интеллекте не только сервис, доступный по запросу, но и научную и инженерную область, в которую можно внести собственный вклад.

Среди местных примеров, в которых я вижу черты такого подхода, — **Misraj AI** и её исследовательская лаборатория **Kawn Lab**.

Эта экосистема заслуживает внимания потому, что не ограничивает себя приложениями и агентами поверх коммерческих моделей. Она занялась исследованием арабских моделей, способов их обучения и специализации, а также созданием данных и контрольных тестов (Benchmarks) — базовых компонентов для любой отрасли, желающей обрести собственную исследовательскую способность.

В апреле 2025 года исследователи Misraj опубликовали работу [*Kuwait 1.5B: An Arabic SLM via Language Injection*](#).

Её исходная идея была исследовательской.

Можно ли внедрить новый язык в малую языковую модель (Small Language Model — SLM), сохранив прежние знания модели и не оплачивая её полное переобучение?

Эксперимент начался с открытой модели TinyLlama. Затем исследователи проверили расширение её слоёв (Layer Extension), добавление арабской лексики в токенизатор (Tokenizer), замораживание частей исходной модели (Layer Freezing) и обучение новых слоёв ради более эффективной языковой адаптации.

В словарь добавили около 26 тысяч арабских токенов (Arabic Tokens). Исследователи изучили влияние расположения новых слоёв, состава обучающих данных и альтернативных архитектурных решений.

Согласно их результатам, среднее улучшение на использованных арабских тестах составило около 8%, при этом прежние способности модели сохранились.

Для меня важна не сама процентная величина.

Перед нами команда, которая рассматривает проблему обучения и межъязыкового переноса знаний (Cross-Lingual Knowledge Transfer), вмешивается в архитектуру модели (Model Architecture) и механизмы её обучения, чтобы найти решение.

Именно такие вопросы открывают путь к созданию технологической индустрии.

Из той же экосистемы Misraj и Kawn вышли и другие работы, среди них **Baseer** — визуально-языковая модель (Vision-Language Model — VLM), предназначенная для понимания арабских документов и извлечения их структуры.

В работе 2025 года *Baseer: A Vision-Language Model for Arabic Document-to-Markdown OCR* исследователи обратились к трудностям оптического распознавания арабского текста (Optical Character Recognition — OCR): от слитного написания букв до разнообразия шрифтов, диакритики и сложной вёрстки страниц.

За основу взяли предварительно обученную мультимодальную модель, затем адаптировали её на арабских данных и разработали тестовый набор Misraj-DocOCR для понимания документов. На опубликованных тестах модель получила показатель ошибки в словах (Word Error Rate — WER) 0,25, опередив по этой метрике системы сравнения в пределах использованного оценочного набора.

У них есть и работы по арабско-английскому переводу — **Mutarjim**, автоматической расстановке диакритики — **Sadeed**, моделированию диалектов — **Lahjawi**, а также направления, связанные с векторными представлениями (Embeddings), мультимодальными наборами данных (Multimodal Datasets) и исследовательским оцениванием. [12](#) [13](#) [14](#)

Эти работы не означают, что Misraj изобрела каждую фундаментальную модель, которой пользовалась. Но это и не критерий оригинальности, который я защищаю.

Разница в том, что компания не просто переименовывает арендованную способность. Она работает с устройством модели, способом её обучения, специализацией или инструментами оценки — и тем самым создаёт знание и технологию, на которых могут строить другие.

Тот, кто смешивает эти вещи, не понимает природы современных исследований.

FlashAttention не изобретала нейронные сети. PagedAttention не изобретала виртуальную память. ASML не изобретала оптическую физику. И всё же каждая из них внесла технический вклад огромной ценности.

Изобретение — не свидетельство о рождении каждого компонента.

Это вклад, благодаря которому сочетание знаний создаёт новую способность.

Именно этот путь, на мой взгляд, воплощают MisraJ и Kawn Lab — один из передовых местных примеров. Вокруг него должны появляться новые лаборатории и исследовательские компании.

Вопрос шире, чем успех одной арабской модели или отдельного продукта.

Речь идёт о **создании исследовательского слоя, способного производить технологии, на которых будут строиться компании будущего.**

Почему нам нужны те, кто создаёт рынок, а не только заполняет его приложениями

Экономист Йозеф Шумпетер (Joseph Schumpeter) писал об инновации как о силе, перестраивающей экономику посредством новых комбинаций (New Combinations). С его именем связано понятие созидательного разрушения (Creative Destruction). [16](#)

В его понимании экономического развития предпринимательство не сводится к выходу на существующий рынок и борьбе за долю. Инновация может создавать новые продукты, способы производства и рынки, меняя сами правила конкуренции.

Для Саудовской Аравии это различие, на мой взгляд, имеет первостепенное значение.

У нас есть возможность построить множество приложений искусственного интеллекта — и это уже происходит в разных отраслях. Но одного изобилия приложений недостаточно, чтобы создать глубокую экономику знаний.

Мы можем получить сотни компаний, которые конкурируют в предоставлении услуг, опираясь на одни и те же глобальные модели. Они будут различаться интерфейсами, ценами и целевыми отраслями, а базовые знания и центры их развития останутся за пределами этих компаний.

Возможен и ещё более странный парадокс: рынок будет переполнен компаниями, называющими себя компаниями искусственного интеллекта, но при любой попытке развивать ядро используемой технологии им придётся обращаться вовне.

Рынок владеет тысячей интерфейсов к разуму — и не имеет достаточно лабораторий, чтобы этот разум развивать.

Поэтому значение глубоко исследовательских компаний выходит далеко за пределы их непосредственных продуктов.

Когда лаборатория вроде Kawn создаёт специализированную модель, метод обучения, набор данных или оценочный тест, она формирует технический актив не для себя одной.

Она помогает создать инфраструктуру, вокруг которой способна вырасти целая отрасль.

Хороший тест задаёт компаниям ясную техническую цель для соперничества. Специализированный набор данных открывает дорогу новым экспериментам. Небольшая высокоэффективная модель может стать основой десятков продуктов. Более быстрый алгоритм логического вывода способен снизить издержки даже тех компаний, которые не участвовали в его создании.

Так возникают рынки знаний (Knowledge Markets) и проявляются эффекты распространения знаний (Knowledge Spillovers).

Подлинная ценность исследования не исчерпывается выручкой компании, которая его провела. В неё входит всё, что это исследование позволяет построить другим компаниям и организациям.

Когда открытие превращается в инструмент, технологию или модель, доступную для лицензирования (Technology Licensing), само знание становится ценным экспортным товаром.

Именно здесь становится очевидной важность компаний типа Misraј и лабораторий типа Kawn.

Они могут участвовать в создании рынка и отрасли: разрабатывать технологические активы, выращивать исследовательские кадры, открывать новые научные задачи и снабжать прикладные компании технологиями, дающими им реальное отличие.

Нам нужны организации, чей эффект подобен эффекту промышленных лабораторий, создавших вокруг своих открытий целые отрасли. Простого финансирования всё большего числа интерфейсов недостаточно.

Это возвращает меня к экономисту Фридриху Хайеку (Friedrich Hayek) и его идее рассеянного знания (Dispersed Knowledge), а также к философу Майклу Полани (Michael Polanyi) и понятию неявного знания (Tacit Knowledge). [15](#) [17](#)

Не всё техническое знание содержится в научных работах и программном коде. Часть его живёт в опыте исследователей, интуиции инженеров, ошибках экспериментов и умении диагностировать проблему прежде, чем становятся ясны её причины.

Эти качества не покупаются по ежемесячной подписке на облачную платформу.

Они накапливаются внутри исследовательских организаций, для которых работа с неизвестным — повседневное занятие.

Если мы хотим построить индустрию, способную обновлять себя, нам нужны организации, которые сохраняют и развивают это накопление, а не компании, меняющие технологическую идентичность с появлением каждой новой коммерческой модели.

Инвестор, финансирующий видимость, получит рынок видимости

Я не думаю, что проблема заключается в недостатке денег, направленных в искусственный интеллект. Вопрос — в том, что мы спрашиваем, решая, куда пойдут эти средства.

Когда компания выходит на рынок и объявляет, что разработала оригинальную технологию искусственного интеллекта, первый вопрос должен звучать так: в чём именно состоит ваш технический вклад?

Вы разработали модель? Изменили её архитектуру? Изобрели метод обучения, алгоритм или систему логического вывода? Владеете весами модели (Model Weights), специализированными обучающими данными или независимой технической инфраструктурой? Можете провести границу между технологией, которую создали сами, и способностью, которую купили?

Это не вопросы с подвохом. Это вопросы об идентичности самой компании.

Я не вижу причин, по которым отвечать на них было бы неловко, — если только маркетинговый рассказ не превосходит реальное достижение.

Если же компания говорит лишь о количестве интеллектуальных агентов (AI Agents), интеграции бизнес-систем, перечне функций и умении модели писать и анализировать, она описывает свой продукт. Но ещё ничего не говорит о том, что изобрела.

Особенно забавно, когда в некоторых презентациях подробно перечисляют способности GPT, а затем приписывают их компании, проводящей презентацию.

Словно модель, пройдя через программный интерфейс из одной организации в другую, по дороге лишилась научной родословной.

Нам необходимо иное инвестиционное понимание глубинных технологий (Deer Tech).

Исследовательскую компанию нельзя считать компанией-разработчиком ПО, которая просто опаздывает с выпуском продукта. От неё нельзя требовать за считанные месяцы добиться того, для чего некоторым открытиям нужны годы.

Но и компания-оболочка не должна получать высокую исследовательскую оценку лишь за умелое употребление научных терминов в презентации.

Нам нужен капитал, готовый принимать научный риск (Scientific Risk), понимать циклы исследований и разработок (R&D Cycles), отличать интеллектуальную собственность (Intellectual Property) от программной интеграции (Software Integration) и признавать: некоторые из самых ценных технологических активов в момент появления ещё не готовы стать массовым продуктом.

Ещё важнее проложить для исследователей ясный путь из лаборатории на рынок.

Если саудовский исследователь создаст алгоритм, заметно ускоряющий логический вывод, архитектуру, которая эффективнее расходует память, или метод, позволяющий небольшим арабским моделям конкурировать в специализированной задаче, его единственным выбором не должны оставаться работа в глобальной компании либо хранение открытия в архиве статей, где его никто не превратит в индустрию.

Он должен найти инвестора, который его понимает, промышленную компанию, которой он нужен, и рынок, готовый его вознаградить.

Нам не нужно больше основателей, умеющих убедить инвестора, будто они изобрели искусственный интеллект. Нам нужно больше исследователей, способных изобрести нечто, после чего такие убеждения станут излишними.

Что, по моему мнению, произойдёт к 2030 году

Я полагаю, что будущее индустрии искусственного интеллекта принесёт с собой большой экономический парадокс.

Чем мощнее будут становиться фундаментальные модели общего назначения (General-Purpose Foundation Models), тем доступнее их способности окажутся для разработчиков и компаний и тем менее редкими станут многие функции, которые сегодня продаются как исключительные преимущества.

Создание интеллектуального агента с несколькими инструментами (Multi-Tool Agent), системы извлечения корпоративного знания или помощника, выполняющего цепочку задач, может стать чем-то вроде сборки стандартных компонентов облачной платформы.

В таком случае многие компании утратят часть превосходства, основанного лишь на соединении этих компонентов.

Я ожидаю, что нынешняя волна компаний-оболочек пройдет через жесткий отбор. Одни превратятся в успешные отраслевые компании-разработчики ПО. Другие станут компаниями услуг и интеграции. Третьи лишатся своего преимущества, когда крупные платформы предложат ту же функцию.

Заявление о владении оригинальным искусственным интеллектом не защитит ни одну компанию от этого сдвига.

Одновременно стратегическое значение четырех областей знания будет расти: архитектур моделей (Model Architectures), эффективности обучения и логического вывода (Training and Inference Efficiency), узкой специализации данных и моделей (Domain Specialization), а также вычислительной инфраструктуры, позволяющей запускать интеллект дешевле и надежнее.

Я думаю, что конкуренция будет не просто гонкой за числом параметров (Parameters), а борьбой за соотношение способности, стоимости, энергопотребления, времени отклика и надежности.

Технология, сокращающая расход памяти, может оказаться экономически значимее приложения, которое добавляет новую функцию в интерфейс чата.

Небольшая специализированная модель может приобрести большую стратегическую ценность, чем более крупная, если решает точную промышленную задачу дешевле и способна работать в локальной инфраструктуре организации (On-Premises Deployment).

Алгоритм, дающий тысячам компаний умеренное снижение стоимости логического вывода, в долгосрочной перспективе может оказаться ценнее десятков приложений, зависящих от одной модели.

Именно к такому будущему, на мой взгляд, должна готовиться Саудовская Аравия.

У нас крупная промышленная экономика, накопленный опыт в энергетике, нефтехимии, водном хозяйстве, транспорте и финансовых услугах, а также условия эксплуатации с задачами, которые трудно решить универсальной диалоговой моделью.

Возможности есть в промышленном искусственном интеллекте (Industrial AI), моделях временных рядов (Time-Series Models), обучении на данных датчиков (Sensor Data Learning), логическом выводе с низкой задержкой (Low-Latency Inference), понимании арабских документов, мультимодальных системах (Multimodal Systems) и технологиях эффективного развёртывания на периферийных устройствах (Edge AI).

Эти области способны стать источниками экспортируемых открытий, если мы будем относиться к ним как к исследовательским задачам, а не только как к возможностям создать очередные приложения.

Поэтому я считаю, что компании и лаборатории типа Misraj и Kawn могут стать важным ядром более широкого преобразования.

Не потому, что они будут обладать монополией на инновации или что одна модель изменит всю отрасль. Они воплощают направление, которое нам необходимо умножить: переход от потребления способностей моделей к созданию знаний, определяющих, на что эти модели способны.

Тогда меняется и сама природа устремления.

Вместо вопроса: сколько саудовских компаний предлагают предприятиям решения на основе GPT?

Появляется другой: сколько технологий, созданных саудовскими командами, используют организации и компании за пределами Саудовской Аравии?

Сколько моделей, алгоритмов, инструментов оценки или инфраструктур логического вывода вышли из саудовской лаборатории в глобальную технологическую экосистему?

Сколько стартапов выросло вокруг научного открытия, а не вокруг коммерческой подписки на языковую модель?

Это не декоративные показатели.

Они измеряют наше положение в цепочке создания ценности знания (Knowledge Value Chain).

Мы можем стать одним из крупнейших рынков искусственного интеллекта в регионе, так и не став производителями влиятельных технологий.

Мы можем вложить миллиарды в вычислительную инфраструктуру, не создав нового инженерного знания, в котором нуждаются другие.

Масштаб даёт присутствие. Оригинальное знание даёт влияние.

Научное прозрение не рождается из формы для доказательств

Современный технологический дискурс, как мне кажется, иногда неверно понимает одну вещь: будто всякое видение будущего должно начинаться с полного набора эмпирических доказательств.

Такой подход годится для проверки конкретного утверждения. Но он не даёт полного описания того, как рождаются открытия и строятся стратегии.

В самой науке всегда была роль научной интуиции (Scientific Intuition), прозрения (Insight), абдуктивного вывода — вывода к лучшему объяснению (Abductive Reasoning) — и способности видеть связи ещё до того, как они оформятся в законченную теоретическую модель.

Учёный и философ Майкл Полани (Michael Polanyi) полагал, что часть знания превосходит то, что мы способны выразить правилами и инструкциями, а опыт порождает способность восприятия, которую трудно полностью свести к механическому описанию. [17](#)

История промышленности полна решений, принятых не из статистической уверенности в будущем, а из инженерного или научного видения, ценность которого большинству современников ещё не была очевидна.

Если бы первопроходцы каждой технологии требовали от будущего заранее представить все доказательства, само лидерство утратило бы смысл.

Поэтому моя ставка на рост ценности глубинных технологий — не заявление о том, что будущее уже написано. Это стратегическое прочтение складывающегося экономического направления: общие способности становятся менее редкими, власть платформ расширяется, а ценность трудно воспроизводимого знания растёт.

Хорошее стратегическое прочтение не ждёт, пока направление станет старой новостью, прежде чем решится назвать его.

Но между предвидением будущего отрасли и заявлением, будто ты изобрёл технологию, которая уже существует у других, проходит совершенно ясная граница.

Интуиция — неотъемлемое право учёного, руководителя и предпринимателя.

А сочинение вымышленной научной биографии для продукта, построенного на GPT, — не интуиция. Это подмена маркировки происхождения.

Создавать то, что нужно миру

Я возвращаюсь к ASML — компании, которая начала свой путь в скромном нидерландском сарае.

Она стала мировой силой не потому, что выпускала приложения, работающие на микросхемах, и не потому, что покупала лучшее из созданного другими, а затем перепродавала под новым именем.

Она выбрала более трудное и глубокое место в цепочке создания ценности.

Она выбрала накопление инженерного знания и превращение его в машины, необходимые компаниям, которые создают продукты для миллиардов людей.

Это был не самый быстрый путь к известности или выручке.

Но он дал такой вид влияния, какого не даёт одно лишь изобилие приложений.

Именно этого я хочу для саудовской индустрии искусственного интеллекта.

Я не хочу, чтобы мы измеряли успех числом арабоязычных интерфейсов поверх GPT, количеством агентов с местными именами или компаний, называющих себя «лидерами искусственного интеллекта», хотя всё их интеллектуальное происхождение сводится к коммерческой подписке у одного из разработчиков моделей.

Я хочу видеть саудовские исследовательские компании, создающие модели, алгоритмы, архитектуры, методы обучения и инструменты оценки, благодаря которым мир будет смотреть на нас как на источник технического знания, а не только как на большой рынок его потребления.

Поэтому в Misraj AI и Kawn Lab я вижу важный пример духа, который нам следует развивать: дух лаборатории, которая ставит вопрос прежде, чем рисует интерфейс; исследует устройство технологии прежде, чем даёт ей торговое имя; стремится производить знание, на котором другие компании смогут строить свои продукты.

Успех компании прикладного уровня, достигшей тысяч клиентов, может быть значимым коммерческим достижением.

Но успех лаборатории, разработавшей технологию, необходимую тысячам компаний, способен открыть целый рынок.

Таково расстояние между участием в существующем рынке и вкладом в создание нового.

Нам не нужно воевать с использованием глобальных моделей. Нам нужно покончить с нелепым маскарадом, в котором импортёр технологии надевает халат её изобретателя.

Кто строит приложение поверх Claude, пусть так и говорит: я построил приложение. Кто предлагает агента на базе GPT, пусть гордится своим продуктом. Кто создал прибыльный SaaS, пусть празднует клиентов и выручку.

Но если кто-то претендует на звание компании глубоких технологий (Deer Tech), глубина должна быть в технологии, а не в лексике, которой он описывает свой интерфейс.

Арендованные умы способны строить успешный бизнес, но они не дают арендатору права заявлять, будто он создал сам разум.

Когда я думаю о следующем десятилетии, меня волнует не то, станем ли мы страной с наибольшим числом приложений, вызывающих интеллект мира.

Меня волнует, станем ли мы страной, которая развивает часть знания, необходимого этому интеллекту, чтобы становиться лучше.

Страны поднимаются до положения создателей знания не тогда, когда ставят множество собственных имён на чужих изобретениях. Они поднимаются, когда имя их изобретений появляется в том, что создают другие.

На экране приложения разница между двумя положениями может казаться небольшой.

Но на весах промышленной истории это разница между вечным клиентом будущего и одним из тех, кто это будущее создаёт.

Источники и литература

1. [Our history](#) — ASML.
2. [Light & lasers — Lithography principles](#) — ASML.
3. [Lenses & mirrors — Lithography principles](#) — ASML.
4. [Cymer](#) — ASML.
5. [The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism](#) — The Quarterly Journal of Economics / Oxford Academic.
6. [Platform envelopment](#) — Strategic Management Journal / Wiley.
7. [New tools for building agents](#) — OpenAI.
8. [FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness](#) — NeurIPS.

9. [Efficient Memory Management for Large Language Model Serving with PagedAttention](#) — ACM SOSP 2023.
10. [Kuwain 1.5B: An Arabic SLM via Language Injection](#) — arXiv.
11. [Baseer: A Vision-Language Model for Arabic Document-to-Markdown OCR](#) — arXiv.
12. [Kawn Lab](#) — Misraj AI.
13. [Research and Publications](#) — Kawn.
14. [Lahjawi: Arabic Cross-Dialect Translator](#) — Association for Computational Linguistics.
15. [The Use of Knowledge in Society](#) — Online Library of Liberty.
16. [Capitalism, Socialism, and Democracy](#) — University of Connecticut course text.
17. [Primary Resources: Transcendence And Self-Transcendence](#) — The Polanyi Society.