

Meta, Microsoft, Nvidia, IBM и другие поддерживают модели ИИ с открытыми весами

Источник: AI News Healthcare

Оригинал: <https://www.artificialintelligence-news.com/news/meta-microsoft-nvidia-ibm-others-back-open-weight-ai/>

бизнес-модели

инфраструктура ИИ

открытый исходный код

регулирование

Две дюжины компаний и организаций подписали открытое письмо, призывающее политиков США защитить модели ИИ с открытыми весами (**open-weight AI models**).

Письмо, опубликованное сегодня (PDF), содержит подписи представителей списка, охватывающего как прямых коммерческих конкурентов, так и организации, чьи бизнес-модели практически не пересекаются: Meta, Microsoft, Nvidia, IBM, Dell Technologies, CrowdStrike, Palantir, ServiceNow.

Аргументация письма строится на сравнении движения за свободное программное обеспечение (**open-source software**) 1980-х годов с нынешней борьбой за то, должны ли веса моделей ИИ распространяться свободно или оставаться заблокированными за коммерческими **API** (интерфейсами прикладного программирования).

Модели с открытыми весами (**open-weight models**) — это системы ИИ, в которых обученные параметры публикуются в открытом доступе, чтобы любой желающий мог скачать, изучить, модифицировать и запустить их на собственном оборудовании. Это отличает их от закрытых моделей, таких как

передовые продукты, предлагаемые OpenAI или Anthropic исключительно через доступ к **API**, где базовые веса никогда не покидают инфраструктуру вендора.

Подписанты рассматривают открытые веса как механизм, позволяющий возможностям ИИ выйти за пределы нескольких высококапитализированных лабораторий и проникнуть в то, что в письме называется рабочими процессами «заводов, больниц, ферм, учебных классов и малого бизнеса».

Их аргументация строится по трем направлениям:

- Открытые веса снижают порог входа для стартапов и государственных учреждений, которые не могут позволить себе обучение передовых моделей с нуля или оплату за каждый токен по ценам передовых моделей для выполнения рутинных задач.
- Они увеличивают конкуренцию на всех уровнях — от чипов до облачной инфраструктуры и приложений, что, согласно письму, удерживает цены на низком уровне и предотвращает захват стоимости ограниченным числом поставщиков.
- Открытые веса также дают корпоративным клиентам способ избежать привязки к конкретному поставщику (**vendor lock-in**), поскольку организации, использующие модели с открытыми весами, контролируют свои собственные данные и могут адаптировать модель под внутренние требования, не завися от дорожной карты или ценовых решений одного вендора.

Аргумент безопасности идет вразрез с интуицией

Самый острый раздел письма напрямую касается рисков, и его стоит прочитать внимательно, так как он переворачивает привычное представление о связи открытых моделей и безопасности.

Письмо признает, что как только веса выпущены, они выходят из-под контроля первоначального разработчика. Модифицированные версии становятся трудно отследить или подвергнуть реверс-инжинирингу. Дообученная (**fine-tuned**) или упрощенная версия открытой модели может распространяться с удаленными защитными барьерами, и механизма отзыва такой модели не существует.

Подписанты утверждают, что ответом должен быть не запрет. Их позиция основывается на сравнении с кибербезопасностью: защитники, сталкивающиеся с атакующими, оснащенными ИИ, нуждаются в доступе к моделям сопоставимых возможностей для обнаружения и симуляции угроз, что закрытые системы с ограниченным доступом предоставить не могут.

Они расширяют это до более широкого утверждения о безопасности, аргументируя это тем, что закрытые модели не являются более безопасными по своей сути, поскольку они могут быть взломаны, использованы не по назначению или выйти из строя так, что внешние исследователи не смогут это наблюдать или проверить. В такой интерпретации концентрация передовых возможностей у небольшого числа закрытых провайдеров создает единые точки отказа, а не устраняет их.

Напротив, открытые модели позволяют сторонним исследователям изучать поведение системы, проводить учения в режиме «красной команды» (**red-team exercises**) и выявлять уязвимости силами множества команд, вместо того чтобы полагаться на внутреннее тестирование одного вендора.

Письмо проводит прямую параллель с аргументом «открытый исходный код более безопасен, чем скрытность», который определял десятилетия дискуссий о безопасности программного обеспечения, хотя в нем не приводятся конкретные данные об обнаружении уязвимостей или статистике инцидентов, применимые именно к системам ИИ.

Дистилляция получает специфическую защиту

В письме выделено место для одной техники, ставшей спорной в кругах ИИ: дистилляция (**distillation**), при которой выходные данные одной модели используются для обучения или улучшения второй модели. Это стандартная практика в исследованиях машинного обучения и разработке продуктов, используемая для оценки, валидации и передачи возможностей между моделями разного размера.

Подписанты проводят черту между дистилляцией как законной техникой и тем, что они называют «незаконными попытками извлечь выгоду из закрытых моделей», утверждая, что первая не должна подпадать под ограничения, направленные на вторую.

Это выглядит как прямой ответ на споры, вспыхнувшие после появления китайских моделей, таких как DeepSeek и Kimi, когда несколько американских лабораторий предположили, что конкурирующие модели были обучены путем дистилляции выходных данных из их собственных закрытых систем без разрешения.

Позиция письма: бороться с незаконным присвоением средств через целевые юридические и коммерческие механизмы, а не через тотальные ограничения технологии, от которой зависит вся отрасль.

Что это означает для предстоящей политической борьбы

Письмо не содержит конкретных законодательных или регуляторных предложений. Это документ, определяющий позицию перед ожидаемыми действиями в области политики ИИ в Вашингтоне; он призывает законодателей расширить доступ к вычислительным мощностям для стартапов и исследователей, финансировать общие наборы данных для обучения и системы оценки, а также избегать того, что авторы называют «преждевременными ограничениями» открытых моделей.

Это следует рассматривать не столько как окончательный политический результат, сколько как индикатор того, к чему стремятся крупнейшие поставщики инфраструктуры и чипов в регуляторном диалоге. У таких игроков, как Nvidia, IBM и Dell, есть прямые коммерческие причины желать процветания экосистем с открытыми весами, поскольку более широкий спектр развертываемых моделей стимулирует продажи вычислительных мощностей и услуг, независимо от того, какая лаборатория создала веса.

Группам по закупкам, выбирающим между развертыванием моделей с открытыми весами и закрытых моделей, следует учитывать, что политическая среда, благоприятствующая одному подходу в ущерб другому, остается неопределенной, и любые ограничения на дистилляцию или открытые релизы могут изменить экономику локально развернутого (**self-hosted**) ИИ в течение одного законодательного цикла.

Перевод выполнен: 24.07.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.