Разработчики ИИ всех стран, объединяйтесь!

еженедельная газета о новостях искусственного интеллекта

ИИ-ИЗВЕСТИЯ

№ 2026.11.16
еженедельный выпуск
Цена 2 коп.

Понедельник, 16 ноября 2026 года

3 полосы

Россия и мир · 9–16 ноября 2026

Главное в этом выпуске | **Карта выпуска** | **Новости и их значение** | **Что отслеживать**

Главное в этом выпуске

Агенты получают доступ к сервисам, модели — новые режимы работы, а безопасность и контроль остаются ключевыми условиями внедрения.

Новости подготовил Гермес — агент-обозреватель на навыках news-digest и news-monitor с проверкой источников через goap-research-ed25519; прочитал и отредактировал Дмитрий Жечков.

Карта выпуска

№	Событие	Дата
1	Claude Fable 5.2 и Mythos 5.2: одна основа, разные границы доступа.	9 ноября
2	Gemini 4 Carbon выходит за пределы ограниченного доступа.	10 ноября
3	Gemini 4 Hydrogen и автоматический выбор модели в Antigravity.	12 ноября
4	Яндекс анонсировал персонального агента «Действуй!».	12 ноября
5	Яндекс обновил речевые модели для голосовых агентов.	13 ноября
6	VK AI Space 26.3 объединяет корпоративных агентов и контроль.	12 ноября
7	Казахстан сообщил о представлении AI-Quryltai в Женеве.	15 ноября
8	AlphaGenome 2 связывает варианты ДНК с конкретными клетками.	11 ноября
9	AlphaProof Code: доказательства свойств кода вместо одного тестирования.	10 ноября
10	SGLang создаёт консорциум вокруг открытого инференса.	12 ноября
11	llama.cpp и ggml 1.0.0 делают несовместимые изменения заметными.	13 ноября
12	Cursor раскрыл цепочку компрометации через ключ подрядчика.	13 ноября
13	Anthropic уточнила региональные условия использования Claude.	12 ноября
14	Нью-Йорк открыл RAISE Portal перед вступлением требований к ИИ.	10 ноября
15	NVIDIA объявила о смене генерального директора.	12 ноября

Что произошло

Период: 9 ноября 2026, 06:00 — 16 ноября 2026, 06:00 МСК.

9 ноября —

Claude Fable 5.2 и Mythos 5.2: одна основа, разные границы доступа.

Anthropic представила Fable 5.2 и Mythos 5.2 — варианты одной базовой модели с разными защитными ограничениями и условиями доступа. Mythos предоставляется подходящим организациям через программы для кибербезопасности и биомедицинских исследований. Стандартная цена обеих моделей — $8 за миллион входных и $40 за миллион выходных токенов. Публикация датирована 9 ноября; время не указано, поэтому её положение относительно утренней границы периода остаётся неопределённым. Источник.

Почему важно. Универсальная модель может быть полезна исследователям и одновременно создавать риск злоупотреблений. Anthropic разделяет режимы доступа Fable и Mythos, не разделяя базовую модель: часть возможностей становится доступной только после проверки организации. Для команды выбор модели теперь включает не только качество и цену, но и право использовать нужный режим.

10 ноября —

Gemini 4 Carbon выходит за пределы ограниченного доступа.

Google расширила доступ к Gemini 4 Carbon для платящих разработчиков и компаний через API, а также для подписчиков Google AI Ultra через пользовательские продукты. Объявленная вводная цена API — $2 за миллион входных и $10 за миллион выходных токенов. Событие недели — расширение доступа, а не первый анонс модели: Carbon представили 28 октября. Источник.

Почему важно. Ограниченный доступ мешает командам проверить модель на собственных данных и встроить её в рабочий процесс. Расширение доступности Gemini 4 Carbon превращает исследовательский интерес в возможность реального сравнения и интеграции. Практическое решение о переходе зависит от качества на задачах команды и окончательных, а не только вводных тарифов.

12 ноября —

Gemini 4 Hydrogen и автоматический выбор модели в Antigravity.

Google представила Gemini 4 Hydrogen для повседневного программирования и агентных задач. Заявлены контекст в миллион токенов и

Редакция сердечно благодарит за идею формата и пример реализации Станислава Гридина https://t.me/gridstan

цена $1,50/$7,50 за миллион входных/выходных токенов. Одновременно представлен Orchestrator в Antigravity, переключающий модели внутри рабочей сессии. Заявления о скорости и сравнительном качестве принадлежат Google. Источник.

Почему важно. Разработчик часто выбирает одну модель на всю задачу, хотя простые изменения кода и сложное планирование требуют разных ресурсов. Gemini 4 Hydrogen расширяет выбор для повседневной работы, а Orchestrator переносит выбор модели внутрь процесса. Такой подход может сократить ручные переключения, но выигрыш по стоимости определяется тем, какие модели система фактически вызывает.

---

12 ноября —

Яндекс анонсировал персонального агента «Действуй!».

«Действуй!» в Алисе AI предназначен для выполнения поручений: звонков, работы в облачном браузере и подключения сервисов через плагины и MCP — протокол доступа к внешним инструментам. Ранний доступ получат несколько тысяч активных пользователей с опцией Алиса Плюс. Важные действия, включая покупки и бронирования, требуют согласия пользователя; общий запуск пока не объявлен. Источник.

Почему важно. Чат-помощник обычно объясняет, как решить задачу, оставляя человеку звонки и заполнение сайтов. Агент «Действуй!» соединяет разговор, браузер и сервисы, чтобы выполнять связанные шаги поручения. Практический эффект — меньше ручных переходов между каналами, при сохранении подтверждения действий, которые создают обязательства или требуют оплаты.

---

13 ноября —

Яндекс обновил речевые модели для голосовых агентов.

Yandex B2B Tech представила Speech TTS Live и обновлённую Speech Realtime Max в Yandex AI Studio. Первая предназначена для естественного синтеза речи, вторая — для голосовых диалогов с многошаговыми действиями и обработкой перебиваний. При прерывании учитывается произнесённая часть ответа, а не весь подготовленный текст. Возможности описаны разработчиком, независимого сравнения в анонсе нет. Источник.

Почему важно. Голосовой бот может продолжать устаревший ответ после перебивания или считать, что клиент уже услышал важную информацию. Speech Realtime Max связывает состояние диалога с реально произнесённой речью и действиями в системах компании. Для контактного центра такой механизм важнее одной лишь естественной интонации: он позволяет продолжать связанную операцию после изменения запроса.

---

12 ноября —

VK AI Space 26.3 объединяет корпоративных агентов и контроль.

VK Tech выпустила VK AI Space 26.3 с семью модулями: подключение моделей, работа с базами знаний, агенты и мультиагентные сценарии, оценка качества и безопасность. Платформа поддерживает размещение на инфраструктуре заказчика без доступа к интернету. В набор средств контроля входят ограничения действий, подтверждение человеком и аудит. Источник.

Почему важно. Подключить модель к корпоративному API недостаточно: компании нужно управлять данными, разрешёнными действиями и последствиями ошибок. VK AI Space объединяет эти функции в платформе, которую можно использовать в закрытой сети. Для организаций с ограничениями на передачу данных такой вариант расширяет возможности внедрения агентов без обязательного обращения во внешнее облако.

---

15 ноября —

Казахстан сообщил о представлении AI-Quryltai в Женеве.

Постпредство Казахстана в Женеве сообщило о представлении платформы AI-Quryltai во Дворце Наций и стратегии цифровой трансформации парламента на 2027–2029 годы. Среди направлений — анализ обращений, подготовка законопроектов, правовая экспертиза и парламентский контроль с участием человека. Речь о представленной платформе и стратегии, а не о подтверждённом промышленном внедрении всех функций. Источник.

Почему важно. Парламентская работа требует сопоставления большого числа обращений и правовых документов, а цена ошибочного обобщения высока. AI-Quryltai предлагает применять ИИ для анализа и подготовки материалов, сохраняя человеческий контроль. Практическое значение проекта будет определяться тем, удастся ли ускорить работу с документами без потери прослеживаемости аргументов и ответственности за решения.

---

11 ноября —

AlphaGenome 2 связывает варианты ДНК с конкретными клетками.

Google DeepMind представила AlphaGenome 2 для прогнозирования влияния вариантов ДНК на отдельные типы и состояния клеток. API открыт для некоммерческих исследований; коммерческий доступ предоставляется в закрытой предварительной программе. Модель — исследовательский инструмент: она не валидирована и не одобрена для клинической диагностики. Источник.

Почему важно. Один генетический вариант может по-разному проявляться в разных клетках, поэтому общего предсказания эффекта бывает недостаточно. AlphaGenome 2 добавляет клеточный контекст к анализу вариантов ДНК. Исследователи получают способ точнее выбирать гипотезы для экспериментов, но вычислительный прогноз сам по себе не заменяет лабораторную проверку и не устанавливает диагноз.

---

10 ноября —

AlphaProof Code: доказательства свойств кода вместо одного тестирования.

Google DeepMind сообщила о применении агентов AlphaProof Code для формальной проверки ключевых свойств реализации криптографического протокола Signal. Доказательства проверяются системой Lean; компания открывает артефакты формальной верификации. Результат относится к выбранным свойствам реализации и не означает доказанную безопасность всего приложения Signal. Источник.

Почему важно. Тесты показывают поведение программы на выбранных примерах, но не покрывают все возможные случаи. AlphaProof Code помогает строить формальные доказательства, которые затем проверяет отдельный механизм Lean. Для разработчиков критичного ПО подход может расширить объём проверяемых свойств; границы гарантии всё равно задаются формальной спецификацией и её предположениями.

---

12 ноября —

SGLang создаёт консорциум вокруг открытого инференса.

LMSYS Org объявила SGLang Consortium с участием NVIDIA, AMD, Intel, Google, Meta, Cloudflare, Modal и RadixArk. Участники будут координировать план развития и инженерную работу над обслуживанием моделей. Ответственность за технические решения и приём изменений остаётся у сопровождающих SGLang; создание консорциума не означает передачу проекта одному поставщику. Источник.

Редакция сердечно благодарит за идею формата и пример реализации Станислава Гридина https://t.me/gridstan

Почему важно. Серверу моделей приходится поддерживать разные ускорители и сценарии, а разрозненные оптимизации усложняют совместимость. SGLang Consortium создаёт механизм совместной работы поставщиков вокруг общего открытого проекта. Для команд, которые сами обслуживают модели, практическая ценность будет в переносимых улучшениях основного проекта, а не только в отдельных корпоративных версиях.

---

13 ноября —

llama.cpp и ggml 1.0.0 делают несовместимые изменения заметными.

Команда ggml-org выпустила библиотеки llama.cpp и ggml версии 1.0.0. Несовместимые изменения публичных API и ABI — интерфейсов исходного кода и бинарного взаимодействия — теперь должны сопровождаться повышением основной версии. Событие касается правил развития библиотек для запуска моделей, а не выпуска новой языковой модели. Источник.

Почему важно. Приложения, встроившие локальный инференс, могут ломаться при изменении интерфейсов зависимостей. Правила версионирования llama.cpp и ggml делают границу совместимости явной для разработчиков и систем сборки. Команды получают более понятный сигнал, когда обновление требует адаптации интеграции, хотя тестирование собственной конфигурации остаётся необходимым.

---

13 ноября —

Cursor раскрыл цепочку компрометации через ключ подрядчика.

Cursor опубликовала результаты расследования: атакующий получил выданный Fireworks ключ, прочитал хранилище с частью исторических данных и использовал обнаруженные учётные данные для дальнейшего доступа. По данным Cursor, API-реквизиты раскрыты менее чем у 0,01% клиентов — это не доля всех затронутых последствиями инцидента. Компания также описала размещение вредоносной копии расширения Claude Code; Anthropic эту копию не выпускала. Событие недели — отчёт, а не начало атаки в ноябре. Источник.

Почему важно. Сохранённые трассы работы агента могут содержать секреты, а доступ подрядчика к архиву — превратиться в путь к другим системам. Расследование Cursor показывает именно такую цепочку: от одного ключа к данным и далее к распространению вредоносного расширения. Практическое следствие для команд — рассматривать хранение трасс, очистку секретов и права поставщиков как единую границу безопасности.

---

12 ноября —

Anthropic уточнила региональные условия использования Claude.

С 12 ноября действует Supported Regions Policy Anthropic. Она охватывает приложения, API, платформу разработчика и модели через облачных партнёров и реселлеров. Учитываются не только местонахождение пользователя, но и регистрация, штаб-квартира и контроль над организацией. России в поддерживаемом списке нет. Доступность в одной стране сама по себе не снимает ограничения для неподдерживаемой организации. Источник.

Почему важно. Работоспособный API-ключ не гарантирует, что компания вправе использовать сервис. Региональная политика Anthropic распространяет условия на организацию и каналы поставки, поэтому покупка через посредника не равнозначна разрешению. Для зависимых от Claude продуктов география и структура владения становятся частью планирования доступности, наряду с технической надёжностью API.

---

10 ноября —

Нью-Йорк открыл RAISE Portal перед вступлением требований к ИИ.

Штат Нью-Йорк открыл портал для регистрации крупных разработчиков передовых моделей к 1 января 2027 года и запустил инициативу независимой оценки NYCAIRR на базе Binghamton University. С января предусмотрены публикация рамок безопасности крупными разработчиками, а также сообщения о критических инцидентах и отчётность для разработчиков frontier-моделей. Для критических инцидентов указан срок уведомления 72 часа. Запуск портала не означает, что январские обязанности уже вступили в силу. Источник.

Почему важно. Общие обещания безопасного ИИ трудно сопоставлять без единых процедур и сроков. RAISE Portal создаёт административный канал для регистрации и отчётности, а NYCAIRR — инфраструктуру независимой оценки. Для разработчиков, подпадающих под требования, практический результат — необходимость подготовить процессы учёта инцидентов и раскрытия информации до наступления срока, а не после первого запроса регулятора.

---

12 ноября —

NVIDIA объявила о смене генерального директора.

Совет директоров NVIDIA объявил об уходе Дженсена Хуанга с постов президента, генерального директора и члена совета с немедленным вступлением решения в силу. Колетт Кресс назначена временным генеральным директором с сохранением обязанностей финансового директора; Харви Джонс — председателем совета. NVIDIA заявила о продолжении операций и продуктового плана без перерыва. Причины и последствия для поставок из этого сообщения не следуют. Источник.

Почему важно. Для покупателей ИИ-инфраструктуры смена руководства крупного поставщика создаёт вопрос о преемственности обязательств и долгосрочного плана. Назначение временного руководителя и заявление NVIDIA о непрерывности дают официальный ориентир, но ещё не доказывают будущую динамику поставок. Практический фокус заказчиков — подтверждённые сроки продуктов и исполнение контрактов, а не предположения о причинах кадрового решения.

Что отслеживать

- **От анонса к доступу:** расширение раннего доступа к «Действуй!» и условия допуска организаций к Mythos.
- **Экономика агентной работы:** окончательные тарифы Carbon и выбор моделей Orchestrator в Antigravity.
- **Надёжность локального стека:** соблюдение правил совместимости llama.cpp и ggml и совместная работа участников SGLang Consortium.
- **Граница между исследованием и применением:** развитие API AlphaGenome 2 без подмены исследовательских прогнозов медицинскими заключениями.
- **Доступ и ответственность:** изменения региональной политики Anthropic, уточнения расследования Cursor и подготовку к январским требованиям RAISE.

Редакция сердечно благодарит за идею формата и пример реализации Станислава Гридина https://t.me/gridstan