



Пресс-релиз

Программное обеспечение KIOXIA AiSAQ расширяет возможности генерации с дополненной выборкой информации (RAG) с использованием ИИ с помощью новой версии библиотеки векторного поиска

Новое программное обеспечение с открытым исходным кодом обеспечивает гибкий баланс между емкостью и производительностью в зависимости от потребностей пользователей и окружающей среды

Германия, Дюссельдорф, 3 июля 2025 г. – В постоянном стремлении улучшить удобство использования поиска векторных баз данных в системах генерации с дополненной выборкой информации (RAG) с использованием ИИ путем оптимизации использования твердотельных накопителей (SSD) компания KIOXIA сегодня объявила об обновлении программного обеспечения KIOXIA AiSAQ™ (All-in-Storage ANNS with Product Quantization). В этой новой версии с открытым исходным кодом представлены гибкие элементы управления, позволяющие разработчикам архитектуры систем определять точку равновесия между производительностью поиска и количеством векторов, которые являются противоположными факторами в фиксированной емкости SSD-накопителя в системе. Полученное в результате преимущество позволяет разработчикам архитектуры систем RAG точно настраивать оптимальный баланс конкретных рабочих нагрузок и их требований без внесения каких-либо изменений в оборудование.

[Впервые представленное в январе 2025 года](#) программное обеспечение AiSAQ компании KIOXIA использует новый алгоритм приближенного поиска ближайшего соседа (ANNS), оптимизированный для твердотельных накопителей, и устраняет необходимость хранения индексных данных в DRAM. Технология KIOXIA AiSAQ, позволяющая выполнять векторный

поиск непосредственно на твердотельных накопителях и снижающая требования к памяти хоста, обеспечивает масштабирование векторных баз данных практически без ограничений по емкости DRAM.

В случае фиксированной установленной емкости SSD-накопителя в системе для увеличения производительности поиска (количество запросов в секунду) требуется больше емкости твердотельного накопителя, потребляемой на каждый вектор. В результате уменьшается количество векторов. И наоборот, чтобы максимально увеличить количество векторов, потребление емкости SSD-накопителя на каждый вектор необходимо уменьшить, в результате чего производительность снижается. Оптимальный баланс между этими двумя противоположными условиями варьируется в зависимости от конкретной рабочей нагрузки. Для нахождения оптимального баланса программное обеспечение KIOXIA AiSAQ предлагает гибкие возможности настройки. Данное последнее обновление позволяет администраторам выбирать оптимальный баланс для различных контрастных рабочих нагрузок в системе RAG. Это обновление делает технологию KIOXIA AiSAQ подходящим алгоритмом приблизительного поиска ближайшего соседа (ANNS) на основе твердотельных накопителей не только для приложений RAG, но и для других приложений, требовательных к векторным данным, таких как автономный семантический поиск.

С увеличением спроса на масштабируемые службы искусственного интеллекта твердотельные накопители являются практичной альтернативой DRAM для управления высокой пропускной способностью и низкой задержкой, которые требуются системам RAG. Программное обеспечение KIOXIA AiSAQ эффективно удовлетворяет этим требованиям за счет крупномасштабного генеративного ИИ без ограничений по ресурсам памяти.

Выпустив KIOXIA AiSAQ как программное обеспечение с открытым исходным кодом, компания KIOXIA подтверждает свою приверженность сообществу исследователей в области искусственного интеллекта, продвигая ориентированные на SSD-накопители архитектуры для масштабируемого ИИ.

«Мы всегда стремимся помочь разработчикам и системным архитекторам в доведении до совершенства настроек производительности и емкости новыми инновационными способами, — сказал Аксель Штёрманн (Axel Störmann), вице-президент и главный технический директор по продуктам памяти и SSD в KIOXIA Europe GmbH. — Теперь, с

внедрением последней версии программного обеспечения KIOXIA AiSAQ, конечные пользователи могут полагаться на мощность твердотельных накопителей для гибкого и эффективного создания масштабируемых систем RAG. Открывая исходный код нашей технологии, мы подтверждаем нашу неизменную приверженность сообществу исследователей в области искусственного интеллекта и предоставляем для него решения, которые являются одновременно мощными и доступными».

Загрузите программное обеспечение с открытым исходным кодом KIOXIA AiSAQ по ссылке:

<https://github.com/kioxia-jp/aisaq-diskann>

#

Примечания:

*AiSAQ от KIOXIA: Новый способ размещения индексных данных All-in-Storage ANNS with Product Quantization является торговой маркой KIOXIA.

*Все другие названия компаний, продуктов и услуг могут быть товарными знаками сторонних компаний.

О KIOXIA

Компания KIOXIA — мировой лидер в сфере решений в области памяти, специализирующийся на разработке, производстве и продаже флэш-памяти и твердотельных накопителей (SSD). В апреле 2017 года ее предшественник, компания Toshiba Memory, отделилась от Toshiba Corporation, которая изобрела флэш-память NAND в 1987 году. Компания KIOXIA стремится вдохновлять мир, раскрывая потенциал систем хранения данных и предлагая продукты, услуги и решения, которые предоставляют заказчикам широкий выбор и создают ценность для общества на основе хранения информации. Инновационная технология 3D флэш-памяти компании KIOXIA — BiCS FLASH™ — формирует будущее хранения данных в приложениях с высокой плотностью, включая современные смартфоны, персональные компьютеры, автомобильные системы, центры обработки данных и системы генеративного ИИ.

Посетите [веб-сайт KIOXIA](#)

Контактные данные для публикации:

KIOXIA Europe GmbH, Hansaallee 181, 40549 Düsseldorf, Germany (г. Дюссельдорф, Германия)

Тел.: +49 (0)211 368 77-0

Электронная почта: KIE-support@kioxia.com