Приоритет Коллекция Назначение Когда внедрять

🟥 Обязательная rag\_chunks Основная коллекция для хранения эмбеддингов чанков документов Сразу, на первом этапе

🟥 Обязательная projects (в PostgreSQL) Хранит список проектов, связан с Qdrant через project\_id Сразу, на первом этапе

🟥 Обязательная documents (в PostgreSQL) Хранит метаданные документов (project\_id, имя, дата и т.д.) Сразу, на первом этапе

| 🟨 Желательная | faq\_chunks | Хранение Q&A пар для быстрых, точных ответов без генерации | Этап 2: после запуска основного RAG |

| 🟨 Желательная | user\_queries | Хранение запросов пользователей и их эмбеддингов для анализа/логов | Этап 2: когда нужна статистика |

| 🟦 Опциональная | context\_embeddings | Шаблоны готовых контекстов для ускорения ответа по типовым вопросам | Этап 3: если есть повторяющиеся кейсы |

| 🟦 Опциональная | semantic\_tags | Тематическая маршрутизация и фильтрация (по темам, категориям) | Этап 3: для улучшения релевантности |

🛠️ Как внедрять поэтапно:

🔹 Этап 1 (Must-Have):

rag\_chunks — хранит текстовые чанки и эмбеддинги

projects — даёт фильтрацию по проектам

documents — позволяет связывать чанк с оригинальным документом

🔸 Этап 2 (Should-Have):

faq\_chunks — когда появится пул повторяющихся вопросов

user\_queries — чтобы видеть, что ищут пользователи (и как часто)

🔵 Этап 3 (Nice-to-Have):

context\_embeddings — можно закэшировать ответы и ускорить отклик

semantic\_tags — если запросы нужно направлять в разные области (например, "юридические", "технические")