

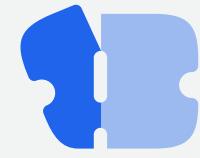


РОСПАРТНЕР

Поиски и управление знаниями
для команд



Проблемы



Разрозненные источники знаний

данные разбросаны по разным сайтам, папкам, документам



Устаревшие материалы

информация быстро теряет актуальность при ручном обновлении



Нет быстрого поиска

сотрудники тратят много времени, пытаясь найти нужные

ИТОГ

Снижение эффективности поддержки и разработки из-за долгого поиска и неактуальных знаний

Целевая аудитория & кейсы



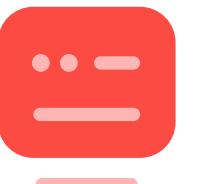
Службы поддержки

оперативно отвечать на вопросы
клиентов с помощью «чата по
документации»



Продуктовые команды

всегда актуальная база знаний
по продукту для внутренних
пользователей



Digital-агентства, веб-разработчики

всегда актуальная база знаний
по продукту для внутренних
пользователей



Внутренние порталы корпораций

обеспечить сотрудников единым
источником правды
и AI-помощником для навигации
по документации

Обзор решения

Мгновенные подсказки (Suggest) с релевантными результатами
<200 мс и глубокий AI-поиск “Ask AI” с генерацией ответа
в реальном времени

Мультипользовательские рабочие
пространства

Автоматический сбор контента

Обработка данных

RAG-генерация (Cloud/On-prem)

Два слоя поиска

Ценность для бизнеса



Единый источник правды

Все знания структурированы, версионированы и доступны командам через единый интерфейс



Авто-актуализация контента

Система сама собирает и обновляет данные, устранивая ручной труд и гарантируя свежесть базы знаний



Готовая AI-цепочка (RAG)

Задействованы генеративные модели, которые отвечают на вопросы, опираясь на проверенные данные компании



Быстрая интеграция

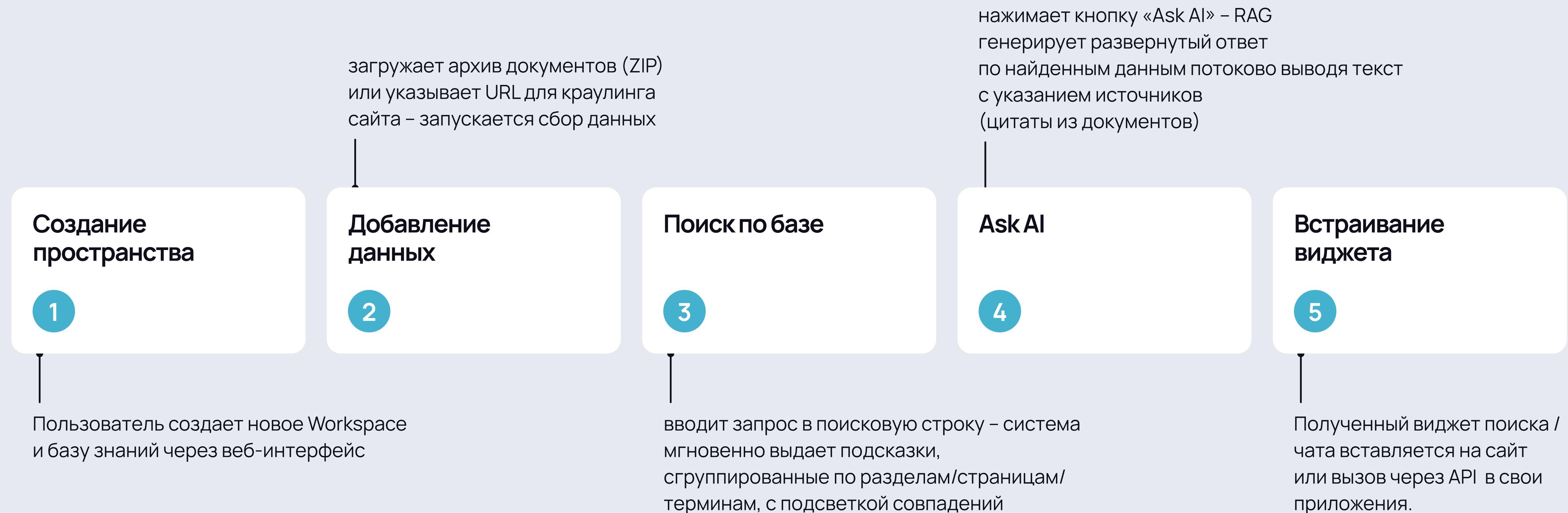
Есть готовый виджет поиска/чата и API – легко встроить в чат-ботов, на сайты и порталы, ускоряя внедрение решения



Контроль доступа и масштабирование

Мультиарендная платформа с ролями пользователей, OAuth и разделением данных по рабочим пространствам повышает безопасность и упрощает масштабирование под новые команды

Скрипт «за 5 шагов»



Масштабирование и отказоустойчивость

Ingestion & чанкинг



Краулинг и импорт

Платформа собирает данные либо автоматически (краулер проходит по ссылкам сайта), либо через ручной импорт файлов



Очистка и нормализация

Текст автоматически очищается от меню, шаблонов, лишнего форматирования – остаётся только содержимое



Нарезка на чанки

Длинные документы дробятся на небольшие фрагменты по $\sim N$ токенов с overlap (перекрытием) для сохранения контекста



Маппинг структуры

Каждый чанк знает, к какому документу и разделу он относится; сохраняется иерархия (раздел \rightarrow документ \rightarrow чанк) плюс версии документов



Фоновые задачи

Весь ingestion выполняется в фоновых очередях с мониторингом – есть статусы задач, автоматические ретраи при ошибках, уведомления об окончании индексации

Два слоя поиска

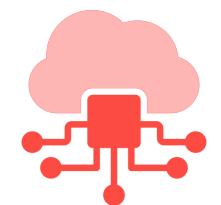
Мгновенные подсказки (/search/suggest)

При вводе запроса система обращается к индексам разделов/страниц/понятий и использует гибридный поиск (объединение BM25 и векторной близости) для выдачи самых релевантных результатов за < 200 мс. Результаты подсвечивают совпадающие слова и показывают источник (раздел/документ)

Генеративный ответ (/chat, RAG)

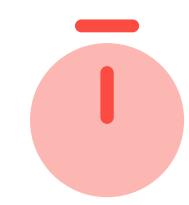
Для глубоких вопросов производится извлечение top-N релевантных чанков, после чего LLM генерирует развернутый ответ на естественном языке. Ответ отдается потоково (как в чатах), сопровождается ссылками на источники. При необходимости можно включить повторное ранжирование результатов перед генерацией для повышения точности

Архитектура



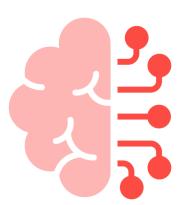
Веб-приложение + API

фронтенд (веб-интерфейс пользователя и виджеты) и бекенд (REST API) работают совместно, обеспечивая управление данными и поиск



Очередь обработки

для ресурсоёмких задач (краулинг, парсинг, эмбеддинг) используется очередь заданий и воркеры, что позволяет обрабатывать контент асинхронно и масштабировать нагрузку



AI интеграция

модели используются для генерации эмбеддингов (векторных представлений чанков) и для LLM-ответов в режиме Ask AI. (В архитектуре предусмотрена возможность подключения собственной LLM при необходимости)



Хранилище данных

сырые документы и структура хранятся в базе знаний, векторные представления – в коллекциях Qdrant (выделенная коллекция на рабочее пространство для изоляции данных)



Мониторинг и логирование

собираются метрики (время ответа, % запросов без результата)

Интеграция & администрирование



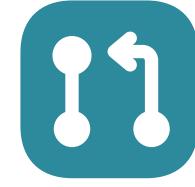
Роли и доступ

Каждое рабочее пространство имеет владельца и ролей (просмотр, редактирование, админ). Интеграция с OAuth-провайдерами (Яндекс, VK и др.) для удобного логина



Безопасность данных

Поддерживаются API-ключи с ограниченными правами (scope), контроль CORS и белый список доменов для виджетов. Все действия пользователей фиксируются (аудит-лог). Регулярно делаются бэкапы снапшоты данных (в т.ч. коллекций векторной базы)



Простая интеграция в продукты

Открытый REST API для поиска и чата; готовые сниппеты и SDK для быстрого встраивания (например, скрипт виджета для сайта на Tilda/Webflow). Есть коллекция запросов Postman и webhooks для расширенных сценариев



Быстрый старт

Подробная документация и примеры позволяют подключить поиск с AI к своим данным за ~10 минут, без необходимости развертывать инфраструктуру с нуля

Дифференциаторы

Русский язык из коробки

Платформа изначально рассчитана на работу с русскоязычными текстами (эмбеддинги и LLM обучены на русском), что выгодно отличает от зарубежных аналогов



No-code импорт и краулинг

Чтобы загрузить данные, не нужен разработчик – материалы автоматически собираются по URL или загружаются файлами, система сама всё обработает



Комбинация классического поиска и AI

Пользователи получают лучшее из двух миров – мгновенный поиск по ключевым словам и интеллигентный AI-помощник, который отвечает на сложные вопросы с опорой на базу знаний

Локальное разёртывание

При необходимости решение может работать полностью в периметре компании – локальная версия LLM и хранение данных в РФ, соответствие требованиям безопасности



Управление знаниями

Не просто поиск – встроено версионирование документов, статусы публикации (черновик, опубликовано и т.п.), закладки и другие инструменты, важные для корпоративной базы знаний



Готовый виджет для сайтов

В отличие от многих конкурентов, мы сразу предоставляем настраиваемый виджет, который легко встроить на сайт без сложной разработки (особенно актуально для конструкторов сайтов вроде Tilda)

Тарифы



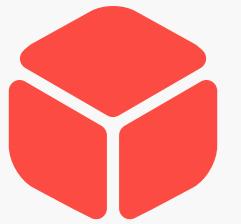
Free

Бесплатно для старта – 1 рабочее пространство, 1 база знаний, ограничение на количество документов и запросов в месяц; на виджете присутствует брендингование



Pro / Team

Подписка для команд – больше рабочих пространств и документов, расширенные возможности (например, более точный поиск с переранжированием, приоритетное обновление краулером), поддержка гостей и ролей для совместной работы



Enterprise

Платное решение для компаний – развёртывание on-premise или в частном облаке, неограниченные объемы данных, интеграция с корпоративными системами (SSO/SCIM), полная изоляция данных, SLA и поддержка, дополнительные услуги (консалтинг по внедрению).



РОСПАРТНЕР - надёжный партнёр в цифровизации бизнеса



18+

Лет опыта работы

50+

Экспертов в команде

100+

Успешных проектов

24/7

Поддержка

Запишитесь на консультацию и демонстрацию решений



Иван Фролов

frolov@hopper-it.ru

Директор по развитию AI-систем

