

AiHummer представляет рекомендации по внедрению ИИ-помощников в бизнес

Разработчик платформы разбирает, как связать ответы ИИ с документами компании, определить допустимые действия и оценить результат пилота. Отдельное внимание уделено выбору инфраструктуры и контролю данных.

Московская компания ООО «АИ ХАММЕР», разработчик платформы AiHummer, подготовила практические рекомендации по внедрению ИИ-помощников в клиентский сервис и внутреннюю работу команд. Материал посвящён переходу от демонстрационного диалога к рабочему процессу, в котором можно проверить исходные данные, выполненное действие и полученный результат.

Компания предлагает начинать внедрение с ограниченной задачи и заранее определять, какие решения система принимает самостоятельно, где требуется подтверждение сотрудника и по каким показателям будет оцениваться пилот. Выбор языковой модели при таком подходе рассматривается вместе с качеством документов, доступными интеграциями и требованиями к эксплуатации.



ИИ-помощник связывает документы компании и обращения клиентов.

Начинать с результата процесса

Задача «автоматизировать общение с клиентами» слишком широка для первого пилота. Более определённый вариант — собрать сведения для расчёта услуги, ответить на вопросы по утверждённым условиям или подготовить обращение к передаче специалисту. Для каждого сценария нужны входные данные, критерий завершения и понятные исключения. Если этих правил нет у сотрудников, подключение ИИ само по себе не устраним неопределённость.

Рабочий результат также стоит отделять от убедительного ответа. Сообщение «заявка создана» ещё не подтверждает, что запись появилась в учётной системе. Проверка должна опираться на результат операции: номер заявки, сохранённый документ или подтверждённую доставку сообщения. Это позволяет оценивать не впечатление от диалога, а выполнение конкретной задачи.

База знаний требует управления

Один из способов связать ответы ИИ с материалами компании — поиск по базе знаний с добавлением найденных фрагментов в запрос к модели. Этот подход называют RAG. В AiHummer он используется для работы с документами и ссылками на источники. Однако само наличие базы знаний не гарантирует, что система найдёт нужный фрагмент, правильно истолкует исключение или выберет действующую редакцию документа.

До запуска полезно определить, какой прайс считается актуальным, кто обновляет инструкции и как выводятся из обращения устаревшие условия. Если в нескольких файлах указаны разные сроки или цены, системе нужен установленный порядок выбора источника. При отсутствии достаточных данных корректным исходом может быть уточняющий вопрос или передача обращения сотруднику, а не попытка заполнить пробел предположением.

Для внутренних помощников отдельно проверяют доступ к документам. Материалы кадровой службы, коммерческие условия и общие инструкции могут быть предназначены для разных сотрудников. Ограничения должны действовать при получении данных, до их передачи модели. Просьба в тексте инструкции «не показывать конфиденциальную информацию» не заменяет техническую проверку прав.



Работа с базой знаний включает поиск источников, контроль актуальности и проверку доступа.

От переписки к проверяемому действию

На примере клининговой компании процесс можно разделить на несколько этапов. Помощник уточняет площадь помещения, вид уборки и дополнительные работы. Затем сведения сопоставляются с прайсом, а предварительный расчёт оформляется в таблицу и передаётся клиенту. Такой сценарий описан в материалах AiHummer. При этом окончательная стоимость подтверждается после осмотра: предварительная смета не должна превращаться в безусловное ценовое обещание.

При проектировании подобных процессов интерпретацию запроса стоит отделять от арифметики и применения тарифных правил. Языковая модель может помочь выделить параметры из сообщения, а расчёт целесообразно выполнять проверяемой функцией или средствами учётной системы. В результате должны быть видны исходные значения и использованная версия прайса. Так сотрудник может проверить, откуда получилась сумма, и исправить исходные данные, не восстанавливая расчёт по переписке.

Сбои интеграций тоже требуют отдельного поведения. Если файл подготовлен, но не доставлен, задача ещё не завершена. Если ответ внешней системы потерян, повторная попытка не должна создавать дубликат заказа. Поэтому в пилот включают проверки не только удачного диалога, но и неполных данных, недоступного сервиса и повторного обращения.



Сценарий клининга: сбор параметров, предварительный расчёт и передача сметы клиенту.

Границы полномочий задаёт приложение

Ответ по справочному документу и изменение записи в рабочей системе имеют разные последствия. На первом этапе помощнику можно предоставить чтение данных и подготовку черновиков, а подтверждение условий заказа от имени компании, удаление записей и другие значимые действия оставить за человеком. Такой порядок позволяет проверить качество работы до расширения полномочий.

В [рекомендациях OWASP](#) по избыточной автономности ИИ-систем отдельно рассматриваются минимально необходимые права, проверка операций и подтверждение человеком действий с существенными последствиями. Практический смысл этих мер — не передавать окончательное решение о доступе самой языковой модели. Команда в сообщении клиента или загруженном файле не должна менять разрешения подключённого инструмента.

Размещение платформы и обработка данных

При выборе между облаком и собственным сервером необходимо разделять два вопроса: где работает приложение и где обрабатывается запрос языковой моделью. Локальная установка платформы может обращаться к внешнему ИИ-сервису. В таком случае запрос и необходимые фрагменты документов передаются его поставщику. Сам термин self-hosted не описывает весь путь данных.

Облачный вариант AiHummer предусматривает выделенный сервер в российском дата-центре и обслуживание инфраструктуры командой разработчика. Альтернатива — установка на собственный Linux-сервер. Платформа поддерживает локальные модели и подключение внешних моделей по ключам заказчика. Конкретный вариант следует выбирать с учётом состава данных, возможностей оборудования и ресурсов команды.

Сравнивать эти варианты только по стоимости лицензии недостаточно. При самостоятельном размещении нужно учитывать администрирование, обновления, резервное копирование и восстановление после сбоев. При использовании внешних моделей — условия обработки данных, доступность провайдера и расходы на запросы. Размещение сервера в определённой стране само по себе не подтверждает выполнение всех требований к обработке информации.



Место установки платформы и способ подключения модели выбирают отдельно.

Каналы связи и история обращений

Подключение нескольких каналов не должно приводить к разным правилам обслуживания. Клиент, который обратился в мессенджере и затем продолжил разговор в другом канале, должен получать согласованные условия. При этом объединять историю разных обращений можно только после надёжного сопоставления пользователя: совпадение отображаемого имени для этого недостаточно.

В AiHummer предусмотрена работа с Telegram, VK, MAX, Авито и другими каналами; набор подключений зависит от тарифа и установленных коннекторов. Для пилота целесообразно выбрать один канал и один процесс, проверить доставку сообщений и передачу обращения сотруднику, а затем расширять охват.



Расширение числа каналов требует согласованных правил и корректной обработки истории обращений.

Оценивать завершённые задачи и полную стоимость

До пилота важно зафиксировать, как выбранная задача выполняется без ИИ: сколько времени она занимает, как часто требует исправлений и сколько обращений остаётся незавершёнными. В тестовый набор стоит включить типичные запросы и исключения: устаревшую цену, неполное описание услуги, противоречивые сведения и просьбу выполнить действие вне полномочий помощника. После изменения документов, модели или интеграции эти проверки повторяют.

Полезные показатели пилота — доля задач, завершённых без исправления сотрудником, время до результата, число ошибок в условиях обслуживания и объём ручной проверки. Передача сложного случая человеку не обязательно означает неудачу: если она предусмотрена правилами, важны полнота собранных сведений и скорость дальнейшей обработки. Общая доля «автоматизированных диалогов» без этих уточнений может скрывать незавершённую работу.

Экономику следует считать по стоимости успешно выполненной задачи с учётом лицензии, инфраструктуры, платных моделей, интеграций и времени сотрудников. Высвобождённые часы не всегда означают прямое сокращение расходов: иногда результатом становится увеличение числа обработанных обращений. Влияние на продажи оценивают отдельно — по сопоставимым периодам или группам, учитывая изменения трафика, цен и сезонности.

Предлагаемый подход не предполагает единого срока окупаемости для всех компаний. Решение о расширении пилота принимают после проверки качества, устойчивости интеграций и фактических затрат. Практическая роль ИИ-помощника определяется тем, какую часть конкретного процесса ему можно доверить и как компания контролирует результат.

Описание платформы и документация доступны на aihummer.ru, демонстрационный интерфейс — на demo.aihummer.ru.

О компании

ООО «АИ ХАММЕР» — московский разработчик платформы AiHummer для работы ИИ-помощников с сообщениями, документами и подключёнными инструментами компании. Продукт доступен в облаке и для установки на собственном сервере.

Контакт для СМИ

Леон Сокиркин, генеральный директор ООО «АИ ХАММЕР»

info@aihummer.ru · +7 (495) 108-01-28

Иллюстрации созданы с помощью ИИ; это концептуальные изображения, не скриншоты продукта.