

# Отчёт о результатах промежуточного тестирования веб-приложения «ИИ-ассистент»

**Неделя:** 4

**Период:** 27.01.2025 – 31.01.2025

**Тестирующая:** Яна Мартынова

---

## 1. Введение

- **Цель тестирования:** Целью тестирования является проверка правильности работы персонального ассистента, созданного для консультирования владельцев собак по вопросам питания. Протестированы параметры настройки ассистента, возможности сбора информации о животном для формирования наиболее релевантного ответа, интеграцию с Google для поиска информации на сайтах производителей кормов и зоомагазинов, интеграцию с Flux Dev для генерации изображений, иллюстрирующих здоровые сбалансированные приемы пищи по системе BARF.
- **Объем тестирования:** Тестирование охватило несколько ключевых аспектов:
  - о Настройка ассистента: роль, личность, стиль общения.
  - о Интеграция с Google для поиска информации и Flux Dev для генерации изображений.
  - о Тестирование параметров модели LLM, производительности и безопасности.
  - о Взаимодействие с пользователем и работа с данными.

## 2. Описание тестируемой системы

- **Тип ассистента:** Ассистент по подбору питания для собак помогает хозяину правильно организовать питание своего питомца. Он консультирует по вопросам выбора типа питания и подбору конкретного корма с учетом особенностей конкретной собаки. Составляет готовые рационы и рассчитывает КБЖУ. Предупреждает об опасности кормления собаки неподходящими продуктами и анализирует ее текущий рацион.
- **Платформа и инструменты:**
  - о **Google** — для поиска информации по запросам.
  - о **Flux Dev** — для генерации изображений на основе запросов.
  - о **Браузер и Интерфейс настройки** для тестирования UI и взаимодействия с пользователем.

## 3. Процесс тестирования

- **Подготовка тестового окружения:**
  - о Настроены параметры ассистента (роль и личность).
  - о Подключены внешние сервисы для поиска информации и генерации

изображений.

- **Методология тестирования:**

- о Выполнены ручные тесты с использованием реальных запросов.
- о Применены чек-листы для проверки основных функций ассистента (поиск, изображения, настройки).

- **Используемые чек-листы:**

- о Чек-листы по настройке ассистента.
- о Чек-листы по тестированию интеграций (Google, Flux Dev).
- о Чек-листы для тестирования параметров модели LLM.

#### **4. Результаты тестирования**

Протестировано качество ответов llm, скорость работы и количество затраченных токенов. Проанализированы ответы ассистента Google Search и произведена оценка сгенерированных изображений ассистентами Text2Image.

В ходе выполнения составлен 31 тест-кейс из которых:

- 19 – для llm
- 10 – для Google Search
- 2 – для Text2Image

Осуществлен прогон всех тест-кейсов (проверки для DeepSeek выполнены не в полном объеме из-за технических проблем на стороне сервера). Оценка качества ответов, особенности и замечания к работе моделей (при необходимости) отмечены в тест-кейсах в описании фактического результата.

По результатам тестирования заведено 5 багов.

- **Тестирование настройки ассистента:**

- о Все параметры (роль, личность, стиль общения) корректно настроены и отображаются в интерфейсе.
- о При изменении параметров и перезагрузке страницы, измененные параметры сохраняются.

- **Тестирование интеграций:**

- о **Поиск в Google:** Ассистент корректно выполняет запросы, например, Результаты поиска соответствуют запросу.

- о **Генерация изображений через Flux Dev:** Изображения генерируются корректно. Однако, было замечено, что он вставляет одну и ту же сгенерированную картинку (да, она соответствует запросу) в разные диалоги и каждый раз списывает токены.  
Заведен баг 1344 (промпт приложен к ишью)

- **Тестирование параметров модели (LLM):**

- о Изменения параметров (например, температура) влияют на стиль

ответов. При низких значениях Temperature и Top P ассистент при сборе информации о животном задает не все вопросы из промта. Хотя они указаны, как обязательные. Чем выше значения штрафов - тем больше ассистент придерживается алгоритму опроса пользователя и дает более релевантные ответы.

о Ассистент стабильно работает при нагрузке (несколько запросов одновременно).

- **Тестирование взаимодействия с пользователем:**

- о Ассистент некорректно собирает информацию о животном. В промпте четко дана инструкция собирать информацию по шагам, переходить на следующий шаг только при ответе пользователя на вопросы из предыдущего шага.

Заведен баг 1347 (промпт приложен к ишью)

- **Тестирование производительности:**

- о Время отклика на запросы составило в среднем 4-5 секунд, что соответствует нормам.

- о При увеличении нагрузки (10 и более запросов одновременно) время отклика увеличивается до 10 секунд.

НО ответы даются односложные без осуществления поиска в интернете.

## 5. Итоги тестирования

- **Общие результаты:** В целом ассистент работает корректно и дает релевантные ответы с учетом потребностей конкретного животного. Однако его рекомендации были бы более адресными, если бы он опрашивал пользователя согласно инструкции из промпта, а не нарушал последовательность вопросов и задавал все указанные вопросы.

Заведено улучшение по настройке ассистента 782

## 6. Приложения

- **Тестовые сценарии:**

- о Тестирование поиска через Google.

1683, 1690

- о Тестирование генерации изображений через Flux Dev.

1684, 1691

- о Проверка производительности.

1720, 1723, 1712

- о Настройки ассистента.

1675, 1679, 1680, 1682

- о Тестирования LLM.

1695, 1701

- о Тестирование взаимодействия с пользователем.>

1702, 1704, 1705, 1706