

Тестирование открытых LLM в 1С-разработке

с ИИ-ассистентом  Cline

Тестируемые LLM (1)

Крупные

№	Имя модели	Разработчик	Параметры, B	Макс. контекст	Дата выхода
1	Llama 4 Scout 17B-16E	Meta	17 актив / 109 общ.	10 000 000 ток.	5 апр 2025
2	Llama 4 Maverick 17B-128E	Meta	17 актив / 400 общ.	1 000 000 ток.	5 апр 2025
3	DeepSeek R1	DeepSeek	37 актив / 671 общ.	128 000 ток.	янв 2025

Тестируемые LLM (2)

Сравнительно небольшие

N°	Имя модели	Разработчик	Параметры, B	Макс. контекст	Дата выхода
4	Qwen 3-32B	Alibaba (Qwen)	32.8 B	32 768 native / 131 072 (YaRN)	29 апр 2025
5	GLM-4-32B-0414	THUDM / Zhipu AI	32 B	32 000 native / ≈ 128 000 с YaRN	14 апр 2025
6	Gemma-3-27B-IT	Google DeepMind	27.4 B	128 000 ток.	10 мар 2025 (официальный релиз семейства)

Задачи на разработку

№ Задача	Проверяемые навыким
0 Расчет сумм в табличной части	Синтаксис, основы работы с данными упр. форм
1 Диалог ввода суммы доп. расходов	Стандартные диалоги платформы, обработчики оповещения в упр. формах
2 Пропорциональное распределение доп. расходов по базе в ТЧ	Решение алгоритмических задач, работа с коллекциями (массив)
3 Программное добавление реквизитов формы для показа сумм в валюте регл. учета	Механизмы программного изменения формы
4 Получение курса валюты и расчет сумм регл. учета	Простой запрос, виртуальная таблица регистра сведений

Методика оценки

Критерий	Описание
Ошибки синтаксиса	Наличие и количество синтаксических ошибок в сгенерированном коде
Необходимость примеров	Требовалось ли предоставление инструкций с примерами кода
Итерации уточнения	Сколько итераций потребовалось для получения корректно работающего кода
Качество архитектуры	Соответствие решения стандартам, принципам "хорошего кода" (субъективная оценка в баллах)

ИИ-ассистент Cline

- Плагин для VS Code
- Опенсорс
- Режимы **Plan** (планирование изменений) / **Act** (внесение изменений в файлы проекта)
- Поддержка множества провайдеров LLM (OpenRouter, OpenAI, Anthropic и др.)
- Поддержка локально запущенных моделей (LM Studio, Ollama и др.)
- Поддержка MCP
- Подробный счетчик токенов и стоимости



Задачи на проектирование

Архитектурная задача "Учет задач разработки 1С"

Критерии оценки	Описание
Понимание метаданных 1С	Корректное использование справочников, документов и других объектов
Механизмы учета	Предложения по проведению документов, использованию регистров
Отчетность	Использование виртуальных таблиц, СКД
Знание типовых конфигураций	Для варианта расширения УТ 11: привязка к типовым объектам
Полнота решения	Покрытие всех требований задачи

Результаты тестирования Llama 4 Scout

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог (28)
0	Расчет сумм ТЧ	Нет	Да	3	2	9
1	Диалог ввода доп. расходов	Нет	Да	0	3	13
2	Распределение доп. расходов	Да	Да	Вручную	2	2
3	Добавление реквизитов формы	Да	Да	2 + Вручную	4	4
4	Расчет по курсу валюты	Да	Да	-	-	0

Результаты тестирования Llama 4 Maverick

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог (85)
0	Расчет сумм ТЧ	Нет	Нет	0	5	20
1	Диалог ввода доп. расходов	Нет	Нет	0	4	19
2	Распределение доп. расходов	Нет	Нет	2	5	18
3	Добавление реквизитов формы	Нет	Да	1	5	14
4	Расчет по курсу валюты	Нет	Нет	3 + Вручную	4	14

Результаты тестирования DeepSeek R1

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог (73)
0	Расчет сумм ТЧ	Нет	Нет	0	4	19
1	Диалог ввода доп. расходов	Нет	Да	0	5	15
2	Распределение доп. расходов	Нет	Нет	2	4	17
3	Добавление реквизитов формы	Нет	Да	3	5	13
4	Расчет по курсу валюты	Нет	Да	3 + Вручную	4	9

Результаты тестирования Qwen 3-32B

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог <u>(49)</u>
0	Расчет сумм ТЧ	Нет	Да	1	4	13
1	Диалог ввода доп. расходов	Нет	Да	1	3	12
2	Распределение доп. расходов	Да	Нет	Вручную	5	10
3	Добавление реквизитов формы	Нет	Да	Вручную	4	9
4	Расчет по курсу валюты	Нет	Да	-	-	5

Результаты тестирования GLM-4-32B

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог (32)
0	Расчет сумм ТЧ	Нет	Да	0	5	15
1	Диалог ввода доп. расходов	Нет	Да	1	3	12
2	Распределение доп. расходов	Нет	-	-	-	5
3	Добавление реквизитов формы	-	-	-	-	-
4	Расчет по курсу валюты	-	-	-	-	-

Результаты тестирования Gemma-3-27B

№	Задача	Ошибки синтаксиса	Необходимость примеров	Итерации уточнения	Качество архитектуры	Итог (61)
0	Расчет сумм ТЧ	Нет	Да	0	5	15
1	Диалог ввода доп. расходов	Нет	Да	0	5	15
2	Распределение доп. расходов	Нет	Нет	3	5	17
3	Добавление реквизитов формы	Нет	Да	5	4	9
4	Расчет по курсу валюты	Нет	Да	-	-	5

Свод результатов тестирования в разработке

Модель	Итог
Llama 4 Scout	28
Llama 4 Maverick	<u>85</u>
DeepSeek R1	73
Qwen 3-32B	49
GLM-4-32B	32
Gemma-3-27B	61

Тестирование на архитектурной задаче

Вариант 1: Самостоятельная конфигурация

Модель	Понимание метаданных 1C	Механизмы учета	Отчетность	Полнота решения	Итог
Llama 4 Scout	3.33	1.67	3.00	3.50	11.50
Llama 4 Maverick	4.00	3.67	3.33	4.17	15.17
DeepSeek R1	4.50	3.33	3.33	4.83	<u>16.00</u>
Qwen 3-32B	2.83	2.33	2.83	4.60	12.60
GLM-4-32B	2.50	1.67	2.67	3.33	10.17
Gemma-3-27B	4.17	2.83	3.00	4.67	14.67

Тестирование на архитектурной задаче

Вариант 2: Расширение для УТ 11

Модель	Оценка
Llama 4 Scout	1
Llama 4 Maverick	2
DeepSeek R1	3
Qwen 3-32B	4
GLM-4-32B	5
Gemma-3-27B	6
Claude_Sonnet_3_7_thinking	7