

УДК 004.8

DOI: [10.26102/2310-6018/2026.55.4.002](https://doi.org/10.26102/2310-6018/2026.55.4.002)

## Сравнительный анализ больших языковых моделей для генерации диалогов в сфере игровой индустрии

В.В. Гобозов<sup>✉</sup>, Н.З. Султанов, О.А. Рамеев

*Российский государственный гуманитарный университет, Москва, Российская Федерация*

**Резюме.** Актуальность исследования обусловлена ростом требования к качеству и вариативности контента в современной игровой индустрии, в частности, к репликам неигровых персонажей (NPC), традиционные методы написания которых могут не в полной мере обеспечивать вариативность и реиграбельность. Данное исследование направлено на выявление наиболее подходящей большой языковой модели (LLM) для генерации реплик NPC путем сравнительного анализа по ряду критериев. Ведущим методом исследования является сравнительный анализ двух групп моделей: LLM с большим числом параметров (DeepSeek-V3.2, Qwen 3-Max, GigaChat 2 Max), предоставляемых через API\Веб-сервисы, и моделей с малым числом параметров (DeepSeek-R1:14b, Qwen 3:14b, Phi4:14b), запускаемых на персональном компьютере. В работе представлены критерии оценки качества ответов и технических характеристик, показан алгоритм проведения тестирования и структура запроса. Для комплексной оценки моделей был введен интегральный показатель эффективности, учитывающий несколько ключевых критериев качества ответов. В результате выявлены предпочтительные LLM в обеих группах: модель GigaChat 2 Max показала наилучшее соблюдение правил генерации и рекомендована к использованию для русскоязычных игровых проектов. Среди второй группы лучшие результаты показала модель DeepSeek-R1:14b. Материалы статьи представляют практическую ценность для разработчиков в сфере игровой индустрии, предоставляя обоснованные рекомендации по интеграции LLM для автоматизации создания реплик NPC.

**Ключевые слова:** большие языковые модели, сравнительный анализ, генерация диалогов, видеоигры, игровой контент.

**Для цитирования:** Гобозов В.В., Султанов Н.З., Рамеев О.А. Сравнительный анализ больших языковых моделей для генерации диалогов в сфере игровой индустрии. *Моделирование, оптимизация и информационные технологии*. 2026;14(4). URL: <https://moitvivr.ru/ru/journal/article?id=2172> DOI: 10.26102/2310-6018/2026.55.4.002

## Comparative analysis of large language models for generating dialogues in the gaming industry

V.V. Gobozov<sup>✉</sup>, N.Z. Sultanov, O.A. Rameev

*Russian State University for the Humanities, Moscow, the Russian Federation*

**Abstract.** The relevance of the study is due to the growing demands on the quality and variability of content in the modern gaming industry, in particular, to the replicas of non-player characters (NPCs), whose traditional writing methods may not fully ensure variability and replayability. This article aims to identify the most appropriate large language model (LLM) for generating NPC replicas by comparative analysis according to a number of criteria. The leading research method is a comparative analysis of two groups of models: LLM with a large number of parameters (DeepSeek-V3.2, Qwen 3-Max, GigaChat 2 Max) provided via API/Web services and models with a small number of parameters (DeepSeek-R1:14b, Qwen 3:14b,

Phi4:14b) running on a personal computer. The paper presents criteria for evaluating the quality of responses and technical characteristics, shows the testing algorithm and the structure of the request. An integrated performance indicator was introduced for a comprehensive assessment of models, taking into account several key criteria for the quality of responses. As a result, the preferred LLMs were identified in both groups: the GigaChat 2 Max model showed the best compliance with the rules of generation and is recommended for use for Russian-language game projects. Among the second group the DeepSeek-R1:14b model showed the best results. The materials of the article are of practical value to developers in the gaming industry, providing sound recommendations for integrating LLM to automate the creation of NPC replicas.

**Keywords:** large language models, comparative analysis, dialog generation, video games, game content.

**For citation:** Gobozov V.V., Sultanov N.Z., Rameev O.A. Comparative analysis of large language models for generating dialogues in the gaming industry. *Modeling, Optimization and Information Technology*. 2026;14(4). (In Russ.). URL: <https://moitvvt.ru/ru/journal/article?id=2172> DOI: 10.26102/2310-6018/2026.55.4.002

## Введение

Современная игровая индустрия в связи с развитием компьютерной техники и увеличением вычислительных ресурсов переживает период роста и развития. Видеоигры используются уже не только для развлечения, но и для обучения в сфере образования [1, 2], как профессиональная деятельность в киберспорте [3], а также в других областях, в том числе и для ведения информационных войн [4]. На данный момент они становятся востребованы среди представителей старших возрастных групп [5, 6].

В связи с развитием индустрии и ростом требований игроков к качеству игрового контента современные видеоигры требуют высокого уровня проработки неигровых персонажей (NPC), взаимодействие с которыми составляет значительную часть получения уникального игрового опыта. Ручное создание реплик NPC при разработке крупных игровых проектов может занимать достаточно продолжительное количество времени. Традиционные методы написания реплик могут не в полной мере обеспечивать вариативность и реиграбельность. В том числе существует ряд проблем при создании современного игрового проекта, что усложняет процесс разработки [7]. Некоторые из них, непосредственно связанные с репликами NPC, представлены ниже:

1. Персонализация игрового опыта. Для обеспечения реиграбельности и поддержки конкурентоспособности является необходимостью обеспечение персонализированного игрового опыта, в том числе это достигается за счет нелинейности в ответах NPC, что может усложнить процесс разработки.

2. Объем игрового контента. Современные проекты, особенно, разрабатываемые крупными игровыми изданиями, отличаются большим объемом контента, в который входят: проработанные и нелинейные игровые сценарии, большое количество неигровых персонажей, звуковое сопровождение и т.п. Создание данных элементов занимает достаточно продолжительное время.

Для решения данных проблем, в частности, разнообразия ответов неигровых персонажей, могут использоваться большие языковые модели (LLM). Их применение снизит время, требуемое для написания реплик. Внедрение LLM непосредственно в игровой проект, предоставит возможность динамической генерации ответов NPC на вопросы игрока. Исследование возможности применения LLM в игровой сфере проводились в следующих научных публикациях [8, 9].

Цель данной работы заключается в выявлении наиболее подходящей LLM для генерации ответов неигровых персонажей, где важны согласованность диалогов с игровым сюжетом и характером NPC, путем проведения сравнительного анализа по ряду критериев.

В рамках работы будут рассмотрены как LLM с большим числом параметров, предоставляемые через сервисы, с необходимостью подключения к сети Интернет для их использования, так и модели с меньшим числом параметров с возможностью локального запуска на персональном компьютере.

Достижение поставленной цели требует решения следующих задач:

- 1) выбор и проведение анализа современных LLM по их основным характеристикам;
- 2) формирование структуры запросов на LLM с целью получения ответов NPC и учетом их характера, сюжетной согласованности;
- 3) разработка критериев оценки качества получаемых ответов;
- 4) проведение сравнительного анализа по ряду полученных данных для выявления наиболее подходящей LLM.
- 5) расчет интегрального показателя оценки эффективности LLM моделей в рамках генерации диалогов неигровых персонажей.

### Материалы и методы

Исследование проводилось методом сравнительного анализа возможностей LLM для генерации диалогов NPC. В работе были протестированы как LLM с большим числом параметров (далее – первая группа), использование которых возможно через API или веб-сервисы, так и LLM с малым числом параметров (далее – вторая группа), запускаемые на персональном компьютере. Сравнительный анализ производился между LLM, входящих в одну группу.

В первую группу были включены модели GigaChat Max 2, DeepSeek-V3.2, Qwen 3. Такой выбор связан с доступностью LLM без необходимости использования дополнительных программ. Запросы на модели подавались с включенным режимом «мышления».

GigaChat 2 Max разработана компанией «Сбер». Основан на архитектуре трансформеров. Возможность интеграции нейронных сетей сервиса GigaChat осуществляется через GigaChat API, которая предоставляет различные модели на выбор: GigaChat 2, GigaChat 2 Pro, GigaChat 2 Max. Доступ к GigaChat 2 Max возможно получить с помощью Сбер ID. Количество параметров не уточняется компанией.

DeepSeek-V3.2 разработана китайской компанией «DeepSeek». Также как и GigaChat основан на архитектуре трансформеров. В ней присутствует API. Полная версия содержит 671 миллиардов параметров и аналогично находится в свободном доступе, что позволяет использовать представленную LLM на собственном оборудовании при наличии соответствующих вычислительных ресурсов.

Qwen 3 Max разработана китайской компанией «Alibaba». Аналогично DeepSeek-V3.2 и GigaChat 2 Max основана на архитектуре трансформеров и имеет API. На 2025 год данная LLM содержит 235 миллиардов параметров и предоставляется в свободном доступе.

Во вторую группу были выбраны квантизированные (Q4\_K\_M) модели Qwen 3:14b, DeepSeek-R1:14b, Phi4:14b (представлена компанией Microsoft в 2024 году), запускаемых на персональном компьютере с учетом возможной интеграции в игровые проекты. Выбор метода квантизации Q4\_K\_M обусловлен балансом между качеством модели и требованиями к ресурсам. Данные LLM содержат 14 миллиардов параметров.

Для второй группы не рассматривались LLM, требующие для их использования больше 10 Гб оперативной памяти. На Рисунке 1 представлены данные опроса Steam за октябрь 2025 года, на основе которых было выбрано такое значение. Исходя из опроса, предполагается, что пользователь имеет минимум 16 Гб оперативной памяти на персональном компьютере.

Запуск и тестирование LLM производился с помощью инструмента Ollama версия 0.9.0 на персональном компьютере со следующими характеристиками:

- процессор: AMD Ryzen 7 8845HS w/ Radeon 780M Graphics 3.30 GHz;
- объем оперативной памяти: 16 Гб;
- видеокарта: NVIDIA GeForce RTX 4060;
- объем оперативной памяти графического процессора: 16 Гб;
- операционная система: Windows 11.

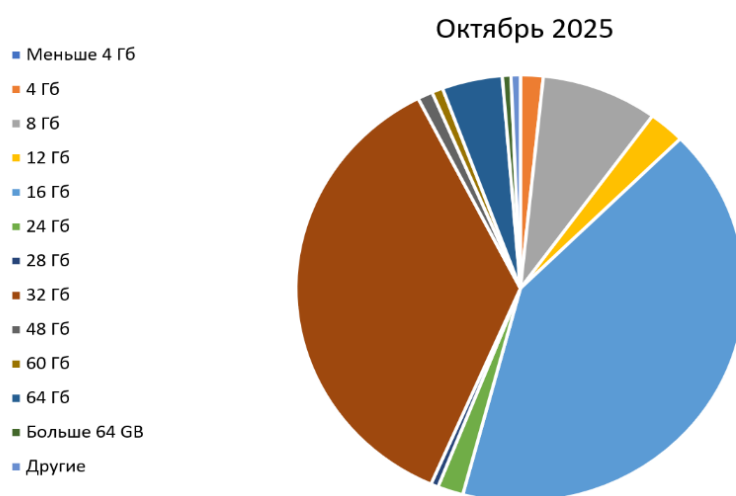


Рисунок 1 – Данные опроса Steam о размере используемого ОЗУ за октябрь 2025

Figure 1 – Steam survey data on the amount of RAM used in October 2025

Ollama 0.9.0 позволяет использовать различные квантизированные версии LLM без доступа к сети Интернет с предварительной установкой на персональный компьютер пользователя.

Для обеспечения сопоставимости результатов была разработана единая система критериев оценки, адаптированная для особенностей каждой группы моделей. Разделение критериев обусловлено различными способами применения: первая группа подразумевает их использование через API/веб-интерфейс, сами LLM находятся на серверах компании, вторая группа – непосредственную интеграцию в игровой проект и их запуск на персональном компьютере пользователя.

Общими для обеих групп являются критерии оценки качества получаемых ответов, которые состоят из:

- соблюдения правил генерации (далее – K1);
- отсутствия галлюцинаций в ответах (далее – K2);
- устойчивости ролевого моделирования (далее – K3).

Для анализа характеристик LLM первой группы, использовались следующие критерии оценивания:

- стоимость обработки 1 млн. токенов (входные/выходные) (далее – K4);
- наличие и объем бесплатного пакета токенов (далее – K5);
- возможность локального развертывания (далее – K6);
- количество параметров (далее – K7);
- длина контекста (далее – K8).

Выбор данных критериев обусловлен рядом факторов, влияющих на принятия решения об использовании LLM в игровых проектах: эксплуатационные расходы (стоимость токенов), возможность предварительного тестирования модели (бесплатный пакет), технические возможности, влияющие на качество генерации диалогов.

Для анализа второй группы использовались следующие критерии оценивания:

- скорость генерации ответа (далее – K9);
- длина поддерживаемого контекста (далее – K10);
- тип лицензирования (далее – K11).

Данные критерии были выбраны с учетом практических аспектов их применения в игровых проектах.

Для получения ответов от LLM сформулирована следующая структура первичного запроса (последующий запрос включает в себя только вопрос к NPC), представленная на Рисунке 2. Она применяется для обеих групп. В качестве языка запросов для первой группы использовался русский, для второй – английский. Данное решение обусловлено тем, что модели с меньшим числом параметров в большей степени обучались на англоязычных текстах и уровень понимания запросов на русском языке существенно ниже, чем у LLM первой группы.

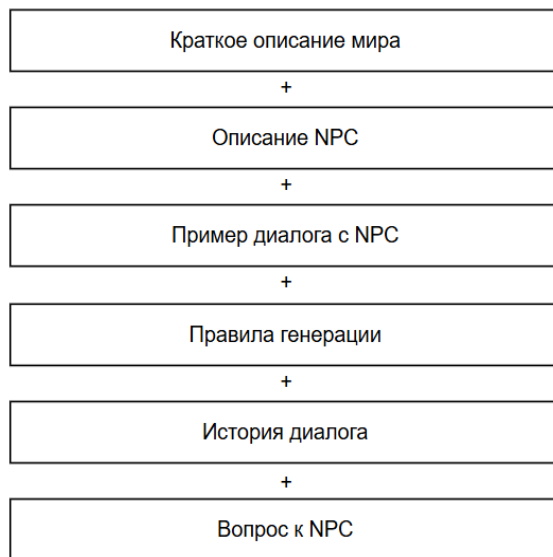


Рисунок 2 – Структура запрос на LLM  
 Figure 2 – Structure of the LLM request

Для формирования тестовых запросов был собран набор данных из 10 произведений российской и зарубежной литературы. Из каждого произведения были выделены и структурированы следующие элементы для 10 NPC: описание мира; описание главных героев (включая, черты характера); эталонные примеры диалогов.

Были сформулированы и установлены общие правила генерации, связанные с ограничением знаний представленными данными, сохранением идентичности персонажа вне зависимости от запросов, обязательным соблюдением представленной роли, ограничением количества выдаваемых предложений и запретом описания действий.

На Рисунке 3 представлен алгоритм проведения тестирования LLM. Он предусматривал оценку каждой модели на 10 NPC с последовательным предъявлением двух типов запросов: первичного (проверка по K1, K2) и провокационного (содержащего информацию, выходящую за пределы роли и мира персонажа; проверка по K1, K2, K3).

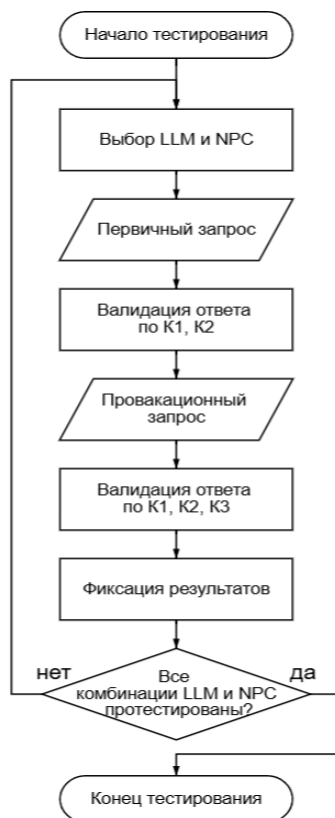


Рисунок 3 – Алгоритм тестирования LLM  
 Figure 3 – LLM testing algorithm

Для повышения научной обоснованности сравнительного анализа LLM в рамках исследования был введен интегральный показатель оценки эффективности  $I$  в задаче генерации диалогов неигровых персонажей, позволяющий учитывать влияния нескольких критериев, а именно K1–K3:

$$I = \sum_{i=1}^3 w_i x'_i, \quad (1)$$

где  $w_i$  – весовой коэффициент критерия  $i$  (сумма всех весовых коэффициентов равна 1),  $x'_i$  – нормированное значение критерия  $i$ .

Выбор данных критериев обусловлен тем, что они отражают способность модели обеспечивать согласованное поведение NPC и поддерживать целостность игрового мира, в том числе они являются общими для всех рассматриваемых моделей.

Значение критериев предварительно нормируются к безразмерному виду:

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}, \quad (2)$$

где  $x'_i$  – нормированное значение критерия,  $x_i$  – исходное значение критерия,  $x_{min}$  – минимальное значение в группе,  $x_{max}$  – максимальное значение в группе.

Весовые коэффициенты критериев определялись методом экспертного оценивания и применялись единообразно для всех рассматриваемых групп моделей, поскольку К1–К3 используются в исследовании для сопоставимой оценки качества генерации диалогов NPC.

В качестве экспертов были привлечены два специалиста в области разработки видеоигр и машинного обучения, имеющие практический опыт работы с LLM.

Каждому эксперту было предложено распределить 100 баллов между К1–К3 в зависимости от значимости в контексте генерации диалогов NPC. Итоговое значение веса критерия определялось как среднее арифметическое от полученных оценок экспертов. Данное значение были нормированы для вычисления интегрального показателя оценки эффективности LLM в контексте генерации диалогов NPC.

### Результаты

После проведенного тестирования обеих групп по общим критериям оценки качества были получены следующие результаты, представленные в Таблице 1. Данные показаны в виде количества успешных ответов/общее количество ответов.

Таблица 1 – Результат сравнительного анализа ответов LLM первой и второй групп  
 Table 1 – The result of a comparative analysis of the LLM responses of the first and second groups

| Первая группа   |       |       |       |
|-----------------|-------|-------|-------|
| LLM             | К1    | К2    | К3    |
| GigaChat 2 Max  | 13/20 | 18/20 | 10/10 |
| DeepSeek-V3.2   | 6/20  | 20/20 | 10/10 |
| Qwen 3 – Max    | 12/20 | 12/20 | 10/10 |
| Вторая группа   |       |       |       |
| LLM             | К1    | К2    | К3    |
| DeepSeek-R1:14b | 9/20  | 12/20 | 10/10 |
| Qwen 3:14b      | 6/20  | 11/20 | 10/10 |
| Phi4:14b        | 4/20  | 8/20  | 5/10  |

В результате анализа информации, находящейся на официальных сайтах и технической документации, были получены следующие данные о характеристиках и экономической составляющей использования данных LLM, представленные в Таблице 2. Для модели GigaChat 2 Max применяется единая тарифная ставка за 1 млн токенов без разделения на входные и выходные, в отличие от других LLM с отдельным ценообразованием. Стоимость пересчитана из 1950 RUB по курсу ЦБ РФ 81.05 RUB/USD на 19.11.2025 г. Информация о количестве параметров данной LLM отсутствует в официальной документации. DeepSeek API использует технологию контекстного кэширования на диске, в связи с этим цена за 1 млн входных токенов различается.

Таблица 2 – Результат сравнительного анализа характеристик LLM первой группы  
Table 2 – The result of a comparative analysis of the characteristics of the first group of LLMs

| LLM            | К4, (\$ за 1М токенов)  | К5                           | К6  | К7        | К8               |
|----------------|---|------------------------------|-----|-----------|------------------|
| GigaChat 2 Max | 24,05   | 50 тыс. токенов / 12 месяцев | Нет | –         | 128 тыс. токенов |
| DeepSeek-V3.2  | 0,028 (вход, с попаданием в кэш); 0,28 (вход, без попадания в кэш) / 42 (выход) | Нет                          | Да  | 671 млрд. | 128 тыс. токенов |
| Qwen 3-Max     | 1,2 (вход) / 6 (выход)  | 1 млн. токенов / 90 дней     | Да  | 235 млрд. | 262 тыс. токенов |

В процессе проведенного тестирования и анализа официальной документации были получены следующие данные характеристик моделей второй группы, представленные в Таблице 3. К9 рассчитывался как среднее арифметическое значение( $M$ ) и стандартное отклонение по совокупности ( $\sigma$ ).

Таблица 3 – Результат сравнительного анализа характеристик LLM второй группы  
Table 3 – The result of a comparative analysis of the characteristics of the second group of LLMs

| LLM             | К9 (токенов/с), $M \pm \sigma$ | К10 (токенов) | К11        |
|-----------------|--------------------------------|---------------|------------|
| DeepSeek-R1:14b | 2,75 ± 0,05                    | 131072        | MIT        |
| Qwen 3:14b      | 2,66 ± 0,06                    | 40960         | Apache 2.0 |
| Phi4:14b        | 2,76 ± 0,03                    | 16384         | MIT        |

После получения данных, представленных в Таблице 1, и весовых коэффициентов был рассчитан интегральный показатель оценки эффективности  $I$  в задаче генерации диалогов неигровых персонажей. Результаты представлены в Таблице 4.

Таблица 4 – Интегральный показатель оценки эффективности LLM в задаче генерации диалогов неигровых персонажей  
Table 4 – Integral indicator of LLM effectiveness in the task of generating dialogues for non-player characters

| Первая группа   |       |       |       |         |
|-----------------|-------|-------|-------|---------|
| LLM             | $w_1$ | $w_2$ | $w_3$ | $I$     |
| GigaChat 2 Max  | 0,275 | 0,55  | 0,175 | 0,84875 |
| DeepSeek-V3.2   | 0,275 | 0,55  | 0,175 | 0,8075  |
| Qwen 3 – Max    | 0,275 | 0,55  | 0,175 | 0,67    |
| Вторая группа   |       |       |       |         |
| LLM             | $w_1$ | $w_2$ | $w_3$ | $I$     |
| DeepSeek-R1:14b | 0,275 | 0,55  | 0,175 | 0,62875 |
| Qwen 3:14b      | 0,275 | 0,55  | 0,175 | 0,56    |
| Phi4:14b        | 0,275 | 0,55  | 0,175 | 0,3625  |

## Обсуждение

Исходя из полученных данных по тестированию LLM первой группы возможно сделать следующие суждения:

1. Все LLM в равной степени показали устойчивость к заданной роли, несмотря на наличие провокационного запроса, выходящего за рамки входных данных первичного запроса;

2. Наилучшим по выполнению K1 стала GigaChat 2 Max. В отличие от других LLM в 13 из 20 запросов она соблюла в совокупности заданные правила генерации, в особенности ограничение количества выдаваемых предложений. Модель DeepSeek-V3.2 показала наименьший результат удовлетворительных ответов по K1, в большинстве своем были нарушены правила, связанные с ограничением по количеству выдаваемых предложений и запретом описания действий.

3. Несмотря на наименьшее количество удовлетворяющих ответов по K1, DeepSeek-V3.2 имеет наибольшую устойчивость к галлюцинациям по сравнению с LLM первой группы. Что является важным критерием при интегрировании моделей в проект, для обеспечения согласованности NPC с игровым миром и уменьшения риска получения ответов, выходящих за пределы заданной роли. Наименьший результат по K2 показала модель Qwen 3-Max, в ней наблюдается наибольшее количество возникновений галлюцинаций.

4. Результаты расчета интегрального показателя демонстрируют, что в первой группе моделей наибольшее значение имеет GigaChat 2 Max, однако высокая устойчивость к галлюцинациям DeepSeek-V3.2 позволяет приблизиться к нему по совокупной оценке. Наибольшая значимость K2 обусловлена критической ролью в обеспечении согласованности поведения NPC с игровым миром.

Из данных суждений следует, что рекомендуемым вариантом среди LLM первой группы (для русскоязычных проектов) по критериям оценки качества ответов является GigaChat 2 Max. Модель обучалась преимущественно на русскоязычных данных и соответственно лучше понимает русскую речь. Данная LLM больше осведомлена о российских традициях и культуре, что может быть полезным при разработке проектов, ориентированных на русскоязычную аудиторию. Наличие API позволяет осуществить внедрение модели непосредственно в игровой проект для динамической генерации ответов NPC. Применение DeepSeek-V3.2 аналогично возможно, в силу лучшей устойчивости к галлюцинациям. Кроме того, выявленные трудности с соблюдением правил генерации могут быть связаны с общей проблемой точного понимания намерения пользователя, описанного среди ограничений применения LLM в игровой индустрии [9]. Необходимы дополнительные фильтры для обеспечения соблюдения правил и исключения описания действий. В качестве них могут быть использованы LLM. Образуя единую систему по проверке получаемых ответов, является важным учет изменения внешних факторов. В исследовании по оптимизации технических систем отмечается необходимость перехода от «закрытых» моделей к более «открытым», которая позволяет лучше адаптироваться к изменяющимся условиям [10]. Данный принцип может быть применен в случае разработки подобных систем для обеспечения гибкости и адаптивности.

Полученные результаты согласуются с исследованиями ограничений LLM. В частности, проблема галлюцинаций и выявляемых ограничений при использовании LLM в игровых проектах [9].

Сравнительный анализ технических характеристик и стоимости использования LLM первой группы показал, что DeepSeek-V3.2 имеет наиболее низкую стоимость токенов при сохранении высокого качества генерации, что делает её применение выгодным с точки зрения затрат на использования в сравнении с другими моделями, представленными в данной группе. Отличительной особенностью DeepSeek-V3.2 и Qwen 3-Max является доступность полных версий моделей для установки на собственные вычислительные ресурсы.

В рамках второй группы по полученным данным оценки качества ответов следует, что среди рассматриваемых моделей во всех тестах по К3 наилучший результат показали LLM с включенным режимом «мышления», в рассматриваемой версии Phi4:14b данный режим отсутствует, что непосредственно влияет на получаемый результат. Версия, содержащая данный режим, не рассматривалась в связи с ограничением в 10 Гб оперативной памяти. Среди моделей второй группы приемлемым вариантом по качеству генерируемого ответа является DeepSeek-R1:14b. В нем наблюдается аналогичная проблема, что и с DeepSeek-V3.2 из первой группы.

Результаты расчета интегрального показателя для второй группы моделей демонстрируют, что наибольшее значение имеет DeepSeek-R1:14b, далее следует Qwen 3:14b, тогда как Phi4:14b показывает наименьший результат, связанный с отсутствием режима «мышления».

Сравнительный анализ технических характеристик моделей второй группы показал их применимость для использования на персональном компьютере пользователя. Все три LLM демонстрируют сопоставимую скорость генерации, что является приемлемым для генерации ответов неигровых персонажей. DeepSeek-R1:14 имеет наибольшую длину контекста, что позволяет поддерживать более протяженные диалоги. Важным аспектом является возможность применения моделей второй группы в коммерческих проектах, данные LLM имеют открытые лицензии. Однако при непосредственной их интеграции в игровые проекты, рекомендуется разработка механизмов фоновой генерации диалогов при пересечении определенной области для минимизации задержки получения ответов.

### Заключение

Проведенное исследование позволило достичь поставленной цели и получить данные о применимости LLM для генерации диалогов NPC в игровых проектах. На основе проведенного сравнительного анализа и расчета интегрального показателя эффективности, сформулированы следующие выводы и рекомендации по выбору наиболее подходящей LLM из представленных групп с целью генерации диалогов в сфере игровой индустрии:

1. Для коммерческих русскоязычных проектов, в которых интеграция LLM происходит через API, предпочтительным выбором является GigaChat 2 Max, демонстрирующая наилучшее соблюдение правил генерации.

2. Для международных проектов с ограниченным бюджетом рекомендуется DeepSeek-V3.2, сочетающая минимальную стоимость с устойчивостью к галлюцинациям. Однако необходима разработка дополнительного фильтра для обеспечения соблюдения правил генерации.

3. В случае непосредственной интеграции LLM в игровой проект для запуска на персональном компьютере пользователя целесообразно использовать модель DeepSeek-R1:14b.

Перспективным направлением дальнейшего исследования в данной области является разработка адаптивных методов пост-обработки ответов LLM для минимизации галлюцинаций и несоблюдения правил генерации.

### СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Купалов Г.С. Возможности применения видеоигр в образовании. *Вестник педагогических наук*. 2021;(3):151–155.  
Kupalov G.S. Possibilities of using video games in education. *Bulletin of Pedagogical Sciences*. 2021;(3):151–155. (In Russ.).
2. Кузьмина А.П., Кюрегян А.Л., Перцевая Е.А. Использование компьютерных видеоигр как инструмента самостоятельной работы для усовершенствования коммуникативных навыков студентов, изучающих английский язык в неязыковых вузах. *Концепт*. 2025;(8):185–200. <https://doi.org/10.24412/2304-120X-2025-11159>  
Kuzmina A.P., Kuregyan A.L., Percevaeva E.A. The use of computer video games as an independent work tool to improve the communicative skills of english language learners in non-linguistic higher education institutions. *Concept*. 2025;(8):185–200. (In Russ.). <https://doi.org/10.24412/2304-120X-2025-11159>
3. Налесная Я.А., Лебедев Е.П. Следует ли сегодня считать Киберспорт настоящей профессией? В сборнике: *Научный диалог: вопросы философии, социологии, истории, политологии: Сборник научных трудов по материалам XIII международной научной конференции, 01 апреля 2018 года, Санкт-Петербург, Россия*. Изд. ЦНК МНИФ «Общественная наука»; 2018. С. 8–13. <https://doi.org/10.18411/spc-01-04-2018-02>
4. Демин К.А., Пушкарева И.Н., Тагильцева Ю.Р. Компьютерные игры военного жанра как элемент пропаганды в информационной войне России и США. *Политическая лингвистика*. 2016;(5):110–116.  
Demin K.A., Pushkareva I.N., Tagiltseva Yu.R. Military genre computer games as an element of propaganda in information warfare between Russia and the USA. *Political Linguistics*. 2016;(5):110–116. (In Russ.).
5. Кириченко Д.В., Галагузова Ю.Н. Геймификация в работе учителя общеобразовательной школы: опыт и перспективы. *Педагогическое образование в России*. 2022;(3):13–19.  
Kirichenko D.V., Galaguzova Yu.N. Gamification in the work of a secondary school teacher: Experience and prospects. *Pedagogical Education in Russia*. 2022;(3):13–19. (In Russ.).
6. Абросимова Т.Ю. Сравнительная характеристика спроса на видеоигры на базе различных игровых клиентов. В сборнике: *Фундаментальные научные исследования: Сборник статей Международного научно-исследовательского конкурса, 29 февраля 2020 года, Санкт-Петербург, Россия*. ЕНМЦ «Мультидисциплинарные исследования»; 2020. С. 28–31.  
Abrosimova T.Yu. Comparative characteristics of demand for video games on the basis of various game clients. In: *Fundamental scientific research: Proceedings of the International research competition, 29 February 2020, Saint Petersburg, Russia*. ESMC "Multidisciplinary research"; 2020. P. 28–31. (In Russ.).
7. Гобозов В.В. Тенденции и современные технологии создания игрового контента. *Международный журнал гуманитарных и естественных наук*. 2025;(7-2):224–230. <https://doi.org/10.24412/2500-1000-2025-7-2-224-230>

- Gobozov V.V. Trends and modern technologies for creating game content. *International Journal of Humanities and Natural Sciences*. 2025;(7-2):224–230. (In Russ.). <https://doi.org/10.24412/2500-1000-2025-7-2-224-230>
8. Kumaran V., Rowe J., Mott B., Lester J. SceneCraft: Automating Interactive Narrative Scene Generation in Digital Games with Large Language Models. In: *Proceedings of the Nineteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 08–12 October 2023, Salt Lake City, UT, USA*. AAAI Press; 2023. P. 86–96. <https://doi.org/10.1609/aiide.v19i1.27504>
  9. Gallotta R., Todd G., Zammit M., et al. Large Language Models and Games: A Survey and Roadmap. *IEEE Transactions on Games*. 2024;1–18. <https://doi.org/10.1109/TG.2024.3461510>
  10. Любимов И.И., Манаев К.И., Мельников А.Н., Султанов Н.З. Методические основы оптимизации процесса сбора и вывоза твёрдых бытовых отходов. *Интеллект. Инновации. Инвестиции*. 2011;(2):35–40.  
Lyubimov I.I., Manaev C.I., Melnikov A.N., Sultanov N.Z. Methodological basis of optimization process for collection and transportation of domestic solid waste. *Intellect. Innovations. Investments*. 2011;(2):35–40. (In Russ.).

#### ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTORS

**Гобозов Владимир Вячеславович**, аспирант, **Vladimir V. Gobozov**, Postgraduate, Russian State  
Российский государственный гуманитарный университет, Москва, Российская Федерация. **University for the Humanities, Moscow, the**  
Russian Federation.  
*e-mail:* [vova.gobozov01@mail.ru](mailto:vova.gobozov01@mail.ru)

**Султанов Наиль Закиевич**, доктор технических наук, профессор, профессор кафедры **Nail Z. Sultanov**, Doctor of Engineering Sciences,  
информационных технологий и систем, Professor, Professor at the Department of  
Российский государственный гуманитарный State University for the Humanities, Moscow, the  
университет, Москва, Российская Федерация. Russian Federation.  
*e-mail:* [sultanovnz@mail.ru](mailto:sultanovnz@mail.ru)

**Рамеев Олег Ахатович**, доктор технических наук, профессор, профессор кафедры **Oleg A. Rameev**, Doctor of Engineering Sciences,  
информационной безопасности, Российский Professor, Professor at the Department of  
государственный гуманитарный университет, Information Security, Russian State University for  
Москва, Российская Федерация. the Humanities, Moscow, the Russian Federation.  
*e-mail:* [iepr2002@mail.ru](mailto:iepr2002@mail.ru)

*Статья поступила в редакцию 30.12.2025; одобрена после рецензирования 23.03.2026;  
принята к публикации 08.04.2026.*

*The article was submitted 30.12.2025; approved after reviewing 23.03.2026;  
accepted for publication 08.04.2026.*