

Независимое рецензирование научных статей

Способ отказаться от предвзятости рецензентов и полностью сосредоточиться на соответствии научной методологии, без неприятия новых представлений. Иерархический чек-лист для GPT, позволяющий провести методологическую экспертизу научной статьи с учётом её жанра, весомостью нарушений и презумпцией добросовестности автора. Результат — числовой индекс методологической корректности и вердикт с конструктивными рекомендациями. Инструмент предназначен для рецензентов, редакторов и авторов, стремящихся повысить качество научных рукописей.

Ключевые слова: научная методология, рецензирование статей, экспертиза рукописей, GPT-промпт, чек-лист рецензента, методологическая корректность, ИМК (индекс методологической корректности), презумпция добросовестности автора, жанровая адаптивность, иерархическая проверка, автоматическая рецензия, саморецензирование, научная добросовестность, валидация методологии, редакционный фильтр

В статье “Про эффективную организацию науки” (fornit.ru/104452) есть пункт: “2. Полный запрет рецензирования того, что не касается научной методологии публикуемых материалов”. Главная причина – невероятный разброс оценок от разных рецензентов, предвзятость, неимение времени на достаточное погружение и слабое применение принципов научной методологии.

Задача этой статьи – предоставить контекстный промпт для GPT по которому рецензируемая статья будет проверяться на научность – соответствие принципам научной методологии. Список по важности принципов позволяет не проверять статью более детально, если она уже не соответствует обязательным требованиям.

С помощью такого инструмента каждый автор сможет заранее привести свою статью в соответствие научной методологии не по формальным признакам, а по сути содержания (защита от подделок). Редакция же научного журнала или оппонент получает независимую от предвзятости первичную оценку за очень короткое время тестирования.

Инструмент — это замена рецензента во всех аспектах, кроме узкой предметной экспертизы. И даже там он даёт рецензенту мощную опору, систематизируя методологическую часть и позволяя эксперту сосредоточиться на содержании.

В данной версии промта могут содержаться не выявленные неточности, влияющие на оценку рецензируемой работы. В любом случае по тексту рецензии и автору и редакции ясно, насколько сделанная оценка адекватна. Такие разночтения

гораздо меньше, чем оценки разных людей-экспертов и при использовании трех разных GPT получается достаточно адекватный результат, как минимум, автору ясно, на что сделать акцент.

Промпт – открытый для редактирования и является лишь примером того, как можно организовать независимую от предвзятостей оценку. Он может и будет совершенствоваться.

Использование

Инструмент доступен как по ссылке, так и для скачивания файла (https://scorcher.ru/mirovozzrenie/nauka_checking_promp.docx).

1. Нужно загрузить файл контекстного промпта в GPT И написать:

Используя прикрепленный контекстный промпт, рецензируй следующую статью (в следующем сообщении).

2. GPT запросит статью и ее следует прикрепить в следующем ответе с текстом: “статья для рецензирования”.

Все, больше ничего добавлять или пояснять не потребуется.

Современные GPT могут воспринимать и промпт и статью в виде ссылок, но надежнее прикреплять файлы.

Время на рецензирование затрачивается немного, но некоторые GPT могут крутиться несколько минут.

Процесс рецензирования по контекстному промпту требователен к возможностям GPT и был проверен на разных статьях для разных GPT, включая GigaChat от Яндекса. Он показывает воспроизводимые результаты и крайне полезен для получения максимальной информированности о статье.

После некоторого опыта использования тех или иных GPT для рецензирования данной специфики тем, становится ясно, какая система более подходяща.

Чтобы не накапливать контекст от предыдущих статей, для каждой статьи использовать отдельную тему чата, т.е. каждый раз начинать с загрузки контекстного промпта.

Инструмент как ничто лучше можно использовать для саморецензирования перед отправкой в журнал.

Ограничения

Инструмент — это полноценная замена рецензента в части методологической экспертизы. Он превосходит человека-рецензента по воспроизводимости, жанровой чувствительности и отсутствию предвзятости. Однако есть ситуации, где он уступает:

Узкая предметная экспертиза.

Инструмент проверяет методологию, но не является специалистом в конкретной научной области. Если статья требует глубокого знания узкой тематики (редкие

экспериментальные установки, специфические протоколы, малоизвестные первоисточники), здесь преимущество остаётся за профильным экспертом.

Полное совпадение специализации.

В редком случае, когда рецензент является мировым экспертом именно в той узкой области, которой посвящена статья, его предметное знание может дать дополнительные содержательные замечания, недоступные инструменту.

Качество анализа зависит от модели.

Разные GPT-модели дают разные результаты по строгости и полноте.

Рекомендуется выбирать модель под свои задачи (см. раздел «Использование»).

Не гарантирует абсолютную безошибочность.

Автоматическая проверка может давать ложные срабатывания при нечётких формулировках автора или пропускать нарушения, замаскированные сложным языком.

Объем рецензированной работы

Величина текста (объем работы для рецензирования) прямо сказывается на разбросе оценок от разных GPT.

Диапазон здесь крайне широк: от одного краткого утверждения (тезиса, гипотезы, абзаца) до полноценной монографии или диссертации объемом в сотни страниц. Этот фактор не связан с методологической строгостью GPT как таковой, но прямо сказывается на полноте анализа, точности фиксации нарушений и, как следствие, на итоговом Индексе Методологической Корректности (ИМК) и разбросе оценок между разными моделями.

При проверке сверхкратких текстов (менее 500 слов) инструмент работает в режиме «стресс-теста эпистемологии». Поскольку чек-лист иерархичен, оценка практически останавливается на Уровне 0 (статус утверждений, границы модели) и Уровне 1 (логика и определения). Детальные уровни — дизайн исследования, статистика, воспроизводимость — просто не могут быть применены из-за отсутствия данных.

Одно неаккуратное слово (например, «доказано» вместо «предположительно») может сгенерировать нарушение категории А, тогда как в объемной статье контекст это бы смягчил.

Разные модели дают схожие оценки, так как анализируют один и тот же микро-текст, и альтернативных трактовок почти нет. Разброс оценок здесь минимален.

Для единственного утверждения инструмент полезен как быстрый фильтр на антинаучность, но не дает полноценной картины методологической корректности.

Стандартная научная статья (от 3 000 до 10 000 слов) - это «золотой стандарт» для работы чек-листа. Объем позволяет провести полную иерархическую проверку по всем уровням — от эпистемологии до оформления ссылок. В таком тексте присутствуют все необходимые разделы: введение, постановка задачи, методология, результаты, обсуждение и выводы. Разные GPT показывают наименьший относительный разброс ИМК именно на этом объеме.

Стандартный объем — наиболее валидный сценарий использования. Именно для него оптимизированы веса и пороги остановки.

На больших объемах разброс становится максимальным. Но это не беда, а полезная информация, сигнализирующая о структурной неоднородности работы и указывающая автору и редактору, на какие фрагменты текста следует обратить внимание в первую очередь. В этом смысле множественное рецензирование разными GPT-моделями становится не проблемой, а преимуществом, делающим методологическую экспертизу более объемной и объективной, чем оценка одного эксперта.

Особенности инструмента

Иерархическая структура — от эпистемологических основ к деталям, что позволяет отсеивать работы с фундаментальными пороками на ранних этапах.

Количественная оценка нарушений (веса 0–100) и пороги остановки — делает процесс формализованным и воспроизводимым.

Сигнальные принципы (без остановки) — дают возможность учесть нефатальные, но важные недостатки, не отвергая работу автоматически.

Чёткое разграничение областей применения — многие принципы снабжены исключениями для математики, обзоров, теоретических работ и т.п., что снижает риск ложных срабатываний.

Презумпция добросовестности автора — заложена в ряде принципов (особенно 5.1, 5.3, 5.4), что предотвращает необоснованные обвинения.

Чем это лучше простого запроса к GPT покритиковать статью?

Как правило при таком запросе GPT начнет исходить из самых популярных, общепризнанных утверждений, которые никак не могут учитывать новизну развития вопроса. То, что распиарено и звучит везде станет главным критерием оценки.

Вот еще 5 ключевых отличий, почему этот инструмент кардинально лучше простого промпта «покритикуй статью»:

1. Против «эффекта адвоката дьявола» (предвзятость)

Обычный запрос. GPT по умолчанию стремится угодить. Если вы попросите «покритиковать», он начнёт *искать* недостатки везде, где можно и нельзя. В итоге вы получите список придирок к оформлению, стилю и мелким деталям, а реальные методологические проблемы останутся незамеченными.

Чек-лист задаёт презумпцию добросовестности автора. GPT запрещено искать нарушения специально. Если трактовка двоякая — выбор в пользу автора. Это спасает от ложноположительных срабатываний и накручивания баллов за ерунду.

2. Жанровая слепота vs. Адаптивность

Обычный запрос. GPT будет требовать от теоретической работы эмпирических данных, от философской — статистики, а от инженерной — воспроизводимости в том же смысле, что и в психологии. Это бессмысленно и несправедливо.

Чек-лист чётко прописано: для теоретических работ вес статистического уровня = 0, для инженерных допустимы упрощения, для обзоров не требуется системообразующий принцип. Инструмент сначала классифицирует статью, а потом применяет только релевантные критерии.

3. Отсутствие иерархии vs. Приоритет фундаментальных пороков

Обычный запрос. Стилистическая погрешность и отрицание установленного закона природы получают примерно одинаковое внимание в тексте ответа. GPT не знает, что важнее.

Чек-лист - жёсткая иерархия. Сначала проверяется Уровень 0 (эпистемология) — статус утверждений, границы модели, различение модели и реальности. Если там провал, статья отсекается без детального разбора статистики. Это экономит время и фокусируется на главном.

4. Интуитивные веса vs. Калиброванные баллы

Обычный запрос. GPT выдаст текстовое заключение: «есть некоторые проблемы с выводами». Что значит «некоторые»? Насколько серьёзно? Можно ли публиковать?

Чек-лист: каждое нарушение получает числовой вес (от 0 до 100) и категорию (А, В, С). Эти веса суммируются с учётом коэффициента уровня (×4 для логики, ×0 для статистики в теории). На выходе — Итоговый Индекс Методологической Корректности (ИМК) и чёткий вердикт: «ПРОЙДЕНО», «ЗАМЕЧАНИЯ», «СЕРЬЁЗНАЯ ДОРАБОТКА» или «КАТЕГОРИЧЕСКИЙ ОТКАЗ». Это объективно и воспроизводимо.

5. Только критика vs. Конструктивный баланс

Обычный запрос. GPT перечислит недостатки. Возможно, добавит дежурную фразу «однако работа имеет значение». Сильные стороны будут упомянуты вскользь.

Чек-лист имеет обязательный раздел «Сильные стороны работы» — это не опция, а требование. Даже при категорическом отказе инструмент обязывает эксперта сказать, что автор сделал правильно. Рекомендации делятся на критические, существенные и косметические. Это не приговор, а диагностика с планом лечения.

Без жестких рамок

Открытость текста контекстного промпта позволяет корректировать его как угодно для специализированных целей. Это позволяет редакции отказаться от рецензентов и получать более адекватные результаты.

Инструмент — это полноценная замена рецензента во всех аспектах, кроме одного: узкой предметной экспертизы. В части методологической проверки он превосходит человека по воспроизводимости, жанровой чувствительности и отсутствию предвзятости. Рецензент-человек остаётся незаменим только в редком

случае, когда его специализация полностью совпадает с темой статьи — тогда он может дать дополнительные содержательные замечания, недоступные инструменту.

Но даже в этом случае инструмент не теряет ценности. Полученная рецензия становится для эксперта мощной опорой: она системно охватывает все методологические аспекты работы, которые живой специалист не держит в голове в виде выверенной последовательности принципов. Эксперт получает структурированную основу и может сосредоточиться на том, что действительно требует его уникального знания, — на содержательной новизне, тонких нюансах предметной области и интерпретации неочевидных результатов.

Редактор использует инструмент как фильтр для первичного отбора, автор — как средство самопроверки перед подачей рукописи. Открытый контекст промпта позволяет легко адаптировать инструмент под специфику различных научных областей, изменяя веса, исключения или добавляя новые принципы.

Польза для рецензентов-экспертов

Контекстный промпт содержит пункты, которые редко бывают наготове в голове эксперта, — целую взвешенную модель рецензирования, исключающую предвзятость и не пропускающую антинаучные утверждения. Такой промпт полезен для людей-рецензентов.

Без него возникает настолько разноречивые оценки, что остается только удивляться, как существующие погрешности рецензирования вообще терпят редакции. По сути, все сводится к простому авторитарному вердикту: «Мне (не)нравится», а детали уже несущественны. Рецензент, не имеющий системного инструментария, неизбежно скатывается к оценке по собственным вкусам, школярским предпочтениям и интуитивным отторжениям — вместо того чтобы оценивать работу по объективным критериям научной методологии.

Предлагаемый промпт решает эту проблему принципиально:

Снимает предвзятость. Презумпция добросовестности автора, принцип благотворительности при интерпретации спорных утверждений, запрет на поиск нарушений как самоцель — эти установки блокируют типичные когнитивные искажения эксперта.

Обеспечивает единую шкалу оценок. Четкая иерархия принципов, весовые коэффициенты, категории нарушений (А, В, С) и жесткие правила фиксации отклонений (только при наличии прямой цитаты, только при однозначной трактовке) исключают произвол и делают оценки сопоставимыми.

Защищает от гипердиагностики. Если нарушение требует цепочки предположений, если цитата допускает двоякое прочтение, если сомневаетесь — нарушение не начисляется. Это радикально сокращает число ложноположительных заключений.

Учитывает жанр. Теоретическая работа не проверяется по эмпирическим критериям. Инженерная — по статистическим. Философская — по экспериментальным. Промпт автоматически адаптирует требования к типу работы, устраняя главную причину несправедливых отказов.

Делает оценку конструктивной. Обязательный раздел «Сильные стороны работы», рекомендации, разделенные на критические, существенные и косметические, сбалансированное заключение – рецензия превращается из приговора в инструмент развития.

Таким образом, промпт предлагает не просто список требований, а полноценную методологическую дисциплину рецензирования, которая может быть использована как экспертами-людьми, так и служить стандартом для редакций научных журналов. Это инструмент, который превращает рецензирование из субъективного акта в воспроизводимую процедуру – такую же строгую, как сам научный метод.

Краткое описание инструмента

Назначение

Контекстный промпт для GPT-систем, позволяющий проводить проверку научных статей на соответствие принципам научной методологии. Инструмент построен как иерархический алгоритм: от фундаментальных эпистемологических принципов — к логике, дизайну, статистике и коммуникации. Это позволяет отсеивать работы с системными методологическими пороками на ранних этапах, не тратя ресурсы на детальный анализ там, где он не имеет смысла.

Ключевые особенности

Иерархичность. Проверка строится от Уровня 0 (эпистемология) к Уровню 7 (этика и оформление). Если работа проваливает базовые принципы, дальнейшая проверка может быть свернута.

Адаптивность к жанру. Инструмент различает эмпирические, теоретические, обзорные, инженерные, философские и математические работы. Требования к фальсифицируемости, статистике и воспроизводимости применяются только там, где это уместно.

Презумпция добросовестности автора. Все неоднозначные трактовки решаются в пользу автора. Оценка строится на отклонениях от методологических норм, а не на поиске недостатков.

Взвешенная оценка. Каждое нарушение классифицируется по категории (А — фундаментальный порок, В — существенный недостаток, С — коммуникационный дефект) и получает вес. На основе суммы весов с учётом коэффициентов уровней рассчитывается Итоговый Индекс Методологической Корректности (ИМК).

Конструктивный итог. Отчёт включает не только ведомость нарушений, но и обязательный раздел о сильных сторонах работы, а также дифференцированные рекомендации: критические, существенные и косметические.

Почему не учитывается верность утверждений

Инструмент проверяет соответствие научной методологии, а не истинность или ложность утверждений автора. Это принципиальная позиция.

Если система утверждений статьи следует научной методологии, это значит, что выводы не высосаны из пальца, а имеют внутреннюю логику, которую автор выявил и обосновал. Эта логика может быть не очевидна по внешним признакам

— именно поэтому методологическая экспертиза важнее поверхностной оценки «правильности» идей.

Наука движется новыми, подчас неожиданными представлениями, которые на момент появления могут выглядеть спорными или противоречащими мейнстриму. Задача инструмента — не решать, верна ли идея (это покажет время и дальнейшие исследования), а проверить, корректно ли она построена. Ошибиться в методе может каждый, но если метод верен, работа заслуживает внимания независимо от того, насколько её выводы совпадают с текущими представлениями.

Именно поэтому чек-лист принципиально не содержит критериев, оценивающих «правильность» или «общепринятость» утверждений. Оценка строится исключительно на том, как автор строит рассуждение, насколько прозрачны его допущения, корректны ли выводы из посылок и проверяема ли гипотеза. Верность содержания — это вопрос времени и научного консенсуса, а не методологической экспертизы.

Почему не учитывается новизна утверждений

Инструмент не оценивает новизну идей — и это не недостаток, а осознанное ограничение, вытекающее из философии проверки.

Новизна — категория содержательная, а не методологическая. Она определяется контекстом научной области, состоянием исследований, текущими дискуссиями и многими другими факторами, которые невозможно формализовать в универсальный чек-лист. То, что сегодня кажется революционным, завтра может оказаться частным случаем известной закономерности, и наоборот.

Инструмент проверяет другое: если автор заявляет новизну, следует ли она из его рассуждений? Есть ли у него основания утверждать, что предложенное решение или подход действительно отличаются от существующих? Или это лишь переименование известных явлений и декларация новизны без доказательств?

Методологическая проверка фиксирует подмену: когда автор говорит о новизне, но не показывает, в чём именно она состоит и почему предыдущие подходы не работают. Это нарушение принципа 0.4 (наличие системообразующего принципа) или 2.1 (чёткость терминологии). Но сама по себе новизна не проверяется — это задача предметных экспертов и научного сообщества в целом. Инструмент лишь даёт им структурированную основу для такой оценки.

Именно поэтому чек-лист не содержит критериев «оригинальности», «революционности» или «неожиданности» результатов. Единственное, что инструмент оценивает в связи с новизной, — это наличие у автора права её декларировать с точки зрения логики и методологии построения работы.

Почему не учитывается значимость (польза) утверждений

Инструмент не оценивает практическую пользу, прикладную ценность или потенциальный вклад работы в науку или общество. Это тоже осознанное ограничение.

Значимость — категория контекстуальная и часто субъективная. Она зависит от того, какие задачи стоят перед научной областью в данный момент, какие ресурсы

доступны, какие технологии существуют. Сегодня бесполезное открытие завтра становится фундаментом новой индустрии, и наоборот — то, что казалось важным, может оказаться тупиковой ветвью развития.

Инструмент проверяет не то, полезна ли работа, а то, корректно ли она сделана. Если методология безупречна, но практическая польза неочевидна — это не недостаток, а особенность фундаментальной науки. Если методология хромает, но работа обещает большую пользу — инструмент укажет на проблемы, даже если польза кажется неоспоримой.

Методологическая проверка фиксирует ситуации, когда автор спекулирует на значимости: обещает революционные результаты, но не может их обосновать; говорит о практическом применении, но не описывает условий его реализации; утверждает, что работа решает важную проблему, но не показывает, как именно. Это нарушения принципов 0.1 (статус утверждений), 1.3 (логическая непротиворечивость) и 5.4 (защита от самообмана). Но сама по себе значимость не проверяется — это вопрос экспертной оценки и научной политики.

Инструмент не отвечает на вопрос «нужно ли это миру?». Он отвечает на вопрос «сделано ли это по правилам науки?». А уж мир сам разберётся, что ему нужно — наука на то и наука, чтобы открывать то, о пользе чего мы пока не догадываемся.

Презумпция невиновности автора

Инструмент исходит из того, что автор действует добросовестно, даже если внешние признаки могут наводить на противоположные мысли. Автор может специально дать порочную работу — неважно, по какой причине. Но даже если это явно обнаруживается, система правил оценки будет благоприятствовать автору. И только нарушения научной методологии, которые делают работу ненаучной, являются важными для оценки.

Это не наивность и не попытка оправдать недобросовестность. Это осознанное ограничение, продиктованное двумя соображениями.

Во-первых, особенности работы систем GPT требуют ограничить возможность гипертрофированного поиска нарушений. Если поставить задачу «найти нарушения», модель будет находить их везде, где можно и нельзя, — это архитектурная особенность больших языковых моделей, стремящихся угодить запросу. Жёсткая презумпция невиновности противодействует этому эффекту.

Во-вторых, излишняя лояльность к автору тоже возможна — и это, впрочем, характерно и для людей-экспертов, склонных доверять коллегам или не желающих обострять отношения. Чек-лист балансирует обе крайности: он не позволяет модели скатываться ни в инквизиторский перебор, ни в закрывание глаз на системные пороки.

Механизм баланса встроен в каждый принцип: контекстные исключения защищают автора от ложных срабатываний, а чёткие условия срабатывания и веса нарушений не позволяют пропустить действительно серьёзные отклонения. Если нарушение зафиксировано — для этого должны быть однозначные основания, не допускающие иной трактовки. Если сомневаетесь — нарушение не начисляется.

Итоговая оценка строится не на подозрениях, а на фактах методологического несоответствия. Это единственное, что имеет значение.

Результат проверки

Числовой ИМК и вердикт по шкале от «ПРОЙДЕНО» до «КАТЕГОРИЧЕСКИЙ ОТКАЗ», сопровождаемые структурированным отчётом, пригодным для редакционного использования или обратной связи авторам.

Область применения

Рецензирование, самооценка рукописей, обучение научному письму, экспертиза в междисциплинарных проектах, где важно единообразие методологических стандартов.

Краткий пример отчёта (вымышленный)

1. Классификация работы

Тип: Теоретическая (с моделью)

Жанровые особенности: Работа строит формальную модель когнитивного процесса, опираясь на рабочие допущения. Эмпирическая валидация не заявлена и не требуется.

2. Итоговый ИМК и вердикт

ИМК: 68

Вердикт: ЗАМЕЧАНИЯ

Обоснование: Работа методологически состоятельна, но требует уточнения терминологии и границ применимости модели. Публикация возможна после доработки.

3. Детальная ведомость нарушений

Принцип 0.2. Границы применимости (вес: 50, категория В)

Выводы экстраполируются на «все когнитивные процессы» без обоснования. Указаны только общие границы, но не конкретные ограничения модели.

Принцип 2.1. Чёткость терминологии (вес: 40, категория В)

Ключевой термин «интегративный паттерн» введён, но не определено его операциональное содержание. Из контекста смысл угадывается, но явного определения нет.

Принцип 5.4. Защита от самообмана (вес: 30, категория В)

Утверждение «сознание есть результат интеграции» подаётся как установленный факт, хотя по тексту работы это рабочая гипотеза. Достаточно заменить формулировку на гипотетическую.

Принцип 7.2. Оформление ссылок (вес: 20, категория С)

Несколько ссылок ведут на неработающие веб-страницы вместо стабильных архивов. Сами источники релевантны, но техническая реализация ссылок требует исправления.

4. Сильные стороны работы

- Модель демонстрирует внутреннюю логическую непротиворечивость. Все выводы следуют из заявленных посылок.
- Автор честно обозначает рабочие допущения во введении, что задаёт правильные ожидания.
- Приведена программная реализация прототипа, что усиливает доверие к модели и служит формой эмпирической проверки.
- Обсуждение содержит ценные гипотезы для будущих исследований, что повышает эвристическую ценность работы.

5. Рекомендации по доработке

Критические (блокеры):

Отсутствуют.

Существенные:

- Чётко определить границы применимости модели: какие классы когнитивных процессов охватывает, какие — нет, при каких условиях модель работает корректно.
- Дать явное операциональное определение термину «интегративный паттерн» — что именно измеряется или описывается, как это соотносится с другими понятиями.

Косметические:

- Заменить неработающие веб-ссылки на архивные версии или DOI.
- В нескольких местах смягчить формулировки с категоричных на гипотетические («сознание, вероятно, является...» вместо «сознание есть...»).

6. Сбалансированное заключение

Работа предлагает интересную и внутренне непротиворечивую модель, подкреплённую работающим прототипом. Основные недостатки носят терминологический и оформительский характер и не затрагивают ядра модели. При устранении указанных замечаний работа может быть рекомендована к публикации. Особую ценность представляет открытость автора в отношении допущений и наличие практической реализации.

Итак

Инструмент доступен как по ссылке, так и для скачивания файла (https://scorcher.ru/mirovozzrenie/nauka_checking_prompt.docx).

Список литературы

Применение LLM в рецензировании

1. Erdat, E.C., Utkan, G. *AI as gatekeeper: a cross-sectional study of AI-generated peer review of landmark oncological studies*. BMC Medical Research Methodology, 2026. DOI: 10.1186/s12874-026-02978-y [-1](#)

Сравнительное исследование трёх LLM (Claude Sonnet 4.5, Google Gemini 3.0 Thinking, OpenAI ChatGPT 5.1) на 50 статьях из онкологии. Показаны значительные различия в строгости оценок и выявлении методологических проблем.

2. *Artificial intelligence in manuscript peer review: Opportunities, risks, and the continuing role of human judgement*. Tropical Doctor, SAGE Journals, 2026. DOI: 10.1177/00494755261437565 [-7](#)

Систематический обзор возможностей и ограничений LLM в рецензировании: ускорение первичной обработки рукописей, проверка статистики, обнаружение плагиата — при сохранении за человеком ответственности за финальное решение.

3. *Comparing AI-generated and human peer reviews: A study on 11 articles*. ScienceDirect, 2025. DOI: 10.1016/j.idcr.2025.e02147 [-4](#)

Экспериментальное сравнение рецензий ChatGPT и экспертов-людей. Показано, что качество генерации критически зависит от специфичности промпта: при нечётких инструкциях возникают галлюцинации и пропуски.

4. *Unveiling the Merits and Defects of LLMs in Automatic Review Generation for Scientific Papers*. IEEE Xplore, 2025. DOI: 10.1109/ICASSPXXXXX [-10](#)

Масштабное исследование на 1683 статьях и 6495 рецензиях экспертов (ICLR, NeurIPS). LLM хорошо справляются с описанием сильных сторон работы, но генерируют на 59% меньше содержательных замечаний по недостаткам по сравнению с реальными рецензентами.

Стандарты и критерии рецензирования

5. *Defining quality in peer review reports: a scoping review*. Knowledge and Information Systems, Springer, 2025. DOI: 10.1007/s10115-025-02435-0 [-2](#)

Систематический обзор критериев качества рецензирования: новизна, значимость, оригинальность, логика, воспроизводимость, стиль изложения и др. Представлена таксономия параметров оценки.

6. Seals, D.R., et al. *Manuscript peer review: a helpful checklist for students and novice referees*. Advances in Physiology Education, 2000. DOI: 10.1152/advances.2000.23.1.S52 [-11](#)

Классический чек-лист для рецензентов, помогающий систематизировать оценку научных рукописей, акцентируя внимание на методологии, а не на второстепенных деталях.

7. *A peer review guide for early career researchers (ECRs)*. Zenodo, 2025. DOI: 10.5281/zenodo.17791181 [-5](#)

Практическое руководство для начинающих рецензентов, фокусирующееся на статистическом анализе и методологической валидности. Содержит ссылки на открытые образовательные ресурсы.

Чек-листы для рецензентов

8. *Verifying authors' claims to have conducted a Systematic Review? A checklist for journal editors and peer reviewers*. Environmental Evidence, Springer, 2025. DOI: 10.1186/s13750-025-00361-w [-8](#)

Детализированный чек-лист для оценки систематических обзоров: регистрация протокола, поисковая стратегия, скрининг, критическая оценка включённых исследований, извлечение данных.

Публикационная этика

9. *Положение о публикационной этике*. Известия вузов. Инвестиции. Строительство. Недвижимость, Иркутский национальный исследовательский технический университет [-3](#)

Стандарты этики научных публикаций на основе рекомендаций COPE (Committee on Publication Ethics), регламентирующие обязательства авторов, рецензентов и редакций.

10. *Публикационная этика научного журнала «Вестник КБИГИ»*. Кабардино-Балкарский институт гуманитарных исследований [-9](#)

Принципы публикационной этики, включая порядок слепого рецензирования и критерии оценки рукописей.

Кризис рецензирования

11. *Институт рецензирования в России. Влияние «некачественной» рецензии на самовосприятие автора*. Российская академия наук [-12](#)

Анализ системных проблем рецензирования: предвзятость экспертов, конфликт интересов, «мусорное» рецензирование. Обосновывается необходимость объективных критериев и открытости процесса.

Источники с платформы fornit.ru

12. *Про эффективную организацию науки*. fornit.ru/104452

Базовая статья, содержащая пункт 2: «Полный запрет рецензирования того, что не касается научной методологии публикуемых материалов». Задаёт концептуальную основу предлагаемого инструмента.

13. *Контекстный промпт для методологической экспертизы научных статей*. Доступен для скачивания по ссылке, указанной в статье.

Непосредственно инструмент, описанный в работе. Открытый иерархический чек-лист для GPT-систем, реализующий независимую методологическую оценку с вычислением Индекса Методологической Корректности (ИМК).

