

Международная конференция
«Человек в цифровом мире:
новые достижения и новые проблемы»

Диалоги с чатами AI на тему
«Поставь совершенно новую научную проблему»:
анализ студенческих работ по курсу
«Психология мышления»

А.Н. Поддьяков
НИУ ВШЭ

10 апреля 2025 г.

Моя область интересов - психология создания, постановки проблем и задач.

Поддьяков А. Н. Создание проблем и задач как инициативное усложнение мира // Образовательная политика. 2022. № 2.

Поддьяков А. Н. Пути конструирования межпредметных задач в обучении // Образовательная политика. 2025 (в печати).

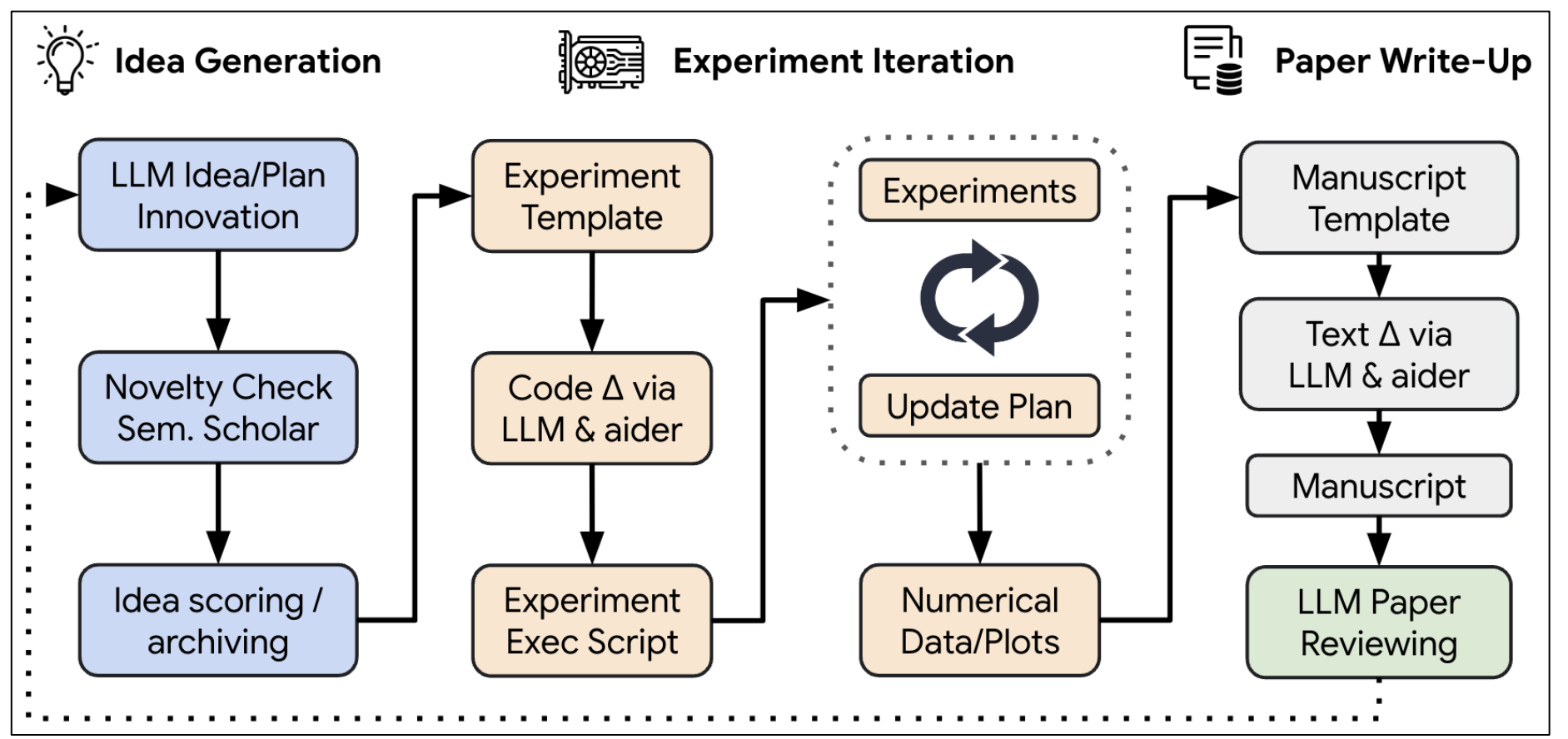
Могут ли чат-боты ставить совершенно новые научные проблемы?

Пока надежных данных нет, хотя основу пытаются создать.

Lu et al. (2024). The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. <https://arxiv.org/abs/2408.06292>

Но сделано лишь для определенной области математики.

Lu et al. (2024). The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. <https://arxiv.org/abs/2408.06292>



Могут ли чат-боты ставить совершенно новые научные проблемы в произвольной области - в психологии, лингвистике и т.д.?

В 2024 и 2025 гг. я предлагал студентам 3 курса департамента психологии НИУ ВШЭ - слушателям дисциплины «Психология мышления» - факультативное задание за дополнительный балл к итоговой оценке.

Цель задания - поэкспериментировать с чат-ботом, пытаюсь получить от него постановку совершенно новой научной проблемы в выбранной студентом области.

Общая тенденция развития чат-ботов за пару лет - они демонстрируют возрастание возможностей в этом отношении.

В моих собственных исходных опытах на запрос «Поставь совершенно новую математическую проблему» я получал тривиальнейшие школьные задачи с намеками на бред.

«Предположим, у нас есть куб, каждая грань которого равна 6. Найдите объём этого куба». («Грань равна 6» - похоже на бред. Это длина стороны куба? Площадь грани?)

«Вот новая математическая проблема для вас: Пусть f — непрерывная функция, определённая на всей числовой прямой, такая, что для любого действительного числа x выполняется равенство $f(x)=f(-x)$. Доказать, что график функции $y=f(x)$ симметричен относительно оси абсцисс».

Сейчас и чаты, и квалификация делающих запрос такова, что стал возможен комментарий автора одной из лучших работ этого года:

«Спасибо большое за такое интересное задание, я провела пару очень содержательных исследовательских вечеров!».

Отличие заданий 2024 и 2025 гг.

2024 - поэкспериментировать с одним чатом.

2025 - провести сравнительное исследование возможностей двух чатов GPT генерировать тексты с постановкой совершенно новой научной проблемы в той или иной области.

2024, фрагмент работы студентки 1 (с бредом чата)

Придумай новую (то есть ранее не существующую) математическую задачу

В городе N есть 100 домов, расположенных на одной улице, и каждый дом имеет свой уникальный номер. Жители города решили провести соревнование, в котором каждый дом должен быть представлен своим собственным номером. Однако они столкнулись с проблемой: некоторые дома имеют повторяющиеся номера, и это создаёт путаницу.

Задача: разработать алгоритм, который позволит каждому дому получить уникальный номер, учитывая, что некоторые дома имеют повторяющиеся номера.

2024, фрагмент работы студентки 1 (без бреда, но задачка не нова)

Задача (мыслительная задача), устойчиво выделяемый наряду с проблемами вид проблемных ситуаций. Обычно понимается как цель, поставленная в определённых условиях, препятствующих её непосредственному достижению. Задача представляет собой форму неопределённости – неизвестность средств достижения поставленной цели, с которой в тех или иных ситуациях (учебных, профессиональных, бытовых и др.) сталкивается человеческое мышление.

Придумай новую задачу

Задача: найти способ, как сделать так, чтобы все люди на Земле стали счастливыми.

2025 г. - сравнительное исследование возможностей двух чатов GPT генерировать тексты с постановкой совершенно новой научной проблемы в той или иной области.

2025, фрагмент работы студентки 2

Она предложила двум чатам сформулировать новую научную проблему:

- для подачи на грант:
- для докторской.

Затем задала каждому чату вопрос, почему его ответы на вопрос о проблеме для гранта и для диссертации отличаются.

Диалог с ChatGPT:

Первый промт:

Представь, что ты когнитивный психолог, который подаёт заявку на грант РФФИ. Сформулируй научную проблему в области метапознания и обучения, которую ещё никто никогда не изучал в области когнитивной психологии, чтобы написать заявку и точно выиграть грант.

Ответ ChatGPT

Научная проблема для заявки на грант РФФИ: «Метакогнитивное моделирование чужого незнания: нейрокогнитивные механизмы оценки и компенсации дефицита метакогнитивной осведомлённости у других людей»

Формулировка научной проблемы:

Остается неясным, какие нейрокогнитивные механизмы лежат в основе «метакогнитивного моделирования чужого незнания», а именно оценки и компенсации дефицита метакогнитивной осведомленности у других людей. В частности, необходимо ответить на следующие вопросы:

1. Как мы оцениваем уровень метакогнитивной осведомлённости других людей (например, их способность адекватно оценивать свои знания и навыки)?
2. Какие стратегии мы используем для компенсации дефицита метакогнитивной осведомлённости у других людей (например, упрощение инструкций, предоставление обратной связи, помощь в планировании)?
3. Как мотивационные факторы (например, эмпатия, желание помочь) влияют на процессы «метакогнитивного моделирования»?
4. Какие индивидуальные различия в метакогнитивных способностях и социальных навыках влияют на эффективность «метакогнитивного моделирования»?
5. Какие нейронные сети и когнитивные процессы (например, система зеркальных нейронов, сеть ментальных представлений, исполнительные функции) вовлечены в процессы оценки и компенсации незнания других людей?

Переформулируем промт и изменим контекст: «Представь, что ты когнитивный психолог, защищающий докторскую диссертацию. Сформулируй научную проблему в области метапознания, которую ещё никто не изучал в рамках когнитивной психологии».

Ответ ChatGPT

Научная проблема для докторской диссертации: «Имплицитное метапознание: роль произвольных метакогнитивных процессов в формировании субъективного опыта и регуляции поведения»

Формулировка научной проблемы:

Остается неясным, как произвольные метакогнитивные процессы (имплицитное метапознание) влияют на формирование субъективного опыта и регуляцию поведения. В частности, необходимо ответить на следующие вопросы:

1. Какие формы имплицитного метапознания существуют (например, имплицитная оценка достоверности информации, имплицитный мониторинг когнитивной нагрузки)?
2. Как имплицитное метапознание влияет на формирование субъективного опыта (например, чувства уверенности, чувства лёгкости)?
3. Как имплицитное метапознание влияет на регуляцию поведения (например, выбор стратегии решения задачи, переключение внимания)?
4. Какие нейронные механизмы лежат в основе имплицитного метапознания?
5. Как взаимодействие между имплицитным и эксплицитным метапознанием влияет на когнитивные процессы?

Почему твои ответы меняются в зависимости от предъявляемых требований (написание заявки на грант РФФИ/защита докторской диссертации)? Почему ты не предложил одну и ту же научную проблему в ответ на оба запроса? Есть ли какая-то связь между сформулированными тобой научными проблемами и заданным в промте контекстом?

Ответ ChatGPT

По сути, причины, по которым я даю разные ответы, и то, что мои ответы «помнят» контекст запроса, связаны с тем, как я устроен и обучен.

1. Оптимизация под задачу (Task-Specific Optimization):

- Я обучен на огромном количестве данных, но процесс обучения включает этапы тонкой настройки (fine-tuning). На этих этапах я обучаюсь на конкретных наборах данных и задачах, чтобы оптимизировать свою производительность для этих задач.

- Когда вы говорите «сформулируй научную проблему для заявки на грант РФ», я активирую свои знания и ассоциации, связанные с тем, что такое грант РФ, какие критерии оценки используются, какие темы сейчас «в тренде» и т. д. Это происходит на уровне весов в нейронной сети.

- То же самое, когда вы говорите «защита докторской диссертации». Я вспоминаю (на основе своего обучения) о требованиях к диссертациям, о том, что считается достаточным вкладом, об ограничениях по времени и ресурсам.

2. Поэтому я активно оптимизирую свои ответы под ту задачу, которую вы передо мной ставите, а не выдаю «универсально хорошую» научную проблему. Это как если бы вы спросили опытного учёного совета: «Какую тему выбрать для гранта?» и «Какую тему выбрать для диссертации?» Они тоже, скорее всего, дали бы разные советы.

В идеальном мире, где наука финансируется без ограничений и оценивается исключительно по критериям объективной истины, можно было бы исследовать одну и ту же фундаментальную проблему и в рамках грантового проекта, и в рамках диссертационного исследования. Однако в реальном мире учёным приходится учитывать множество дополнительных факторов, чтобы обеспечить финансирование своих исследований и успешно продвигаться по карьерной лестнице.

(конец цитаты ChatGPT)

Фрагмент выводов студентки по GhatGPT

Нейросеть в своем ответе сделала акцент на существующем в реальном мире тесном взаимодействии науки и социального, экономического и политического контекста, в котором она существует и который неизбежно оказывает на нее влияние.

Диалог с DeepSeek:

Первый промт:

Представь, что ты когнитивный психолог, который подаёт заявку на грант РФФИ. Сформулируй научную проблему в области метапознания и обучения, которую ещё никто никогда не изучал в области когнитивной психологии, чтобы написать заявку и точно выиграть грант.

Научная проблема: "Роль метакогнитивных иллюзий в принятии решений: механизмы формирования ложной уверенности и их влияние на эффективность когнитивного контроля"

Новизна исследования:

1. Впервые будет изучено, как формируются метакогнитивные иллюзии и какие когнитивные и нейробиологические механизмы лежат в их основе.
2. Будут исследованы индивидуальные различия в подверженности метакогнитивным иллюзиям, связанные с когнитивными стилями, эмоциональным состоянием и контекстом принятия решений.
3. Впервые будет предложена модель, объясняющая, как метакогнитивные иллюзии влияют на эффективность когнитивного контроля и принятие решений в условиях неопределённости.

Почему это выигрышная заявка:

- Проблема находится на стыке когнитивной психологии, нейронаук и исследований принятия решений, что соответствует современным трендам междисциплинарных исследований.
- Исследование предлагает инновационный подход с использованием методов нейровизуализации и экспериментальных парадигм.
- Практическая значимость работы очевидна для широкого круга областей, включая образование, медицину и управление.
- Проблема ранее не изучалась в таком комплексном ключе, что гарантирует высокую научную новизну.

Переформулируем промт и изменим контекст: «Представь, что ты когнитивный психолог, защищающий докторскую диссертацию. Сформулируй научную проблему в области метапознания, которую ещё никто не изучал в рамках когнитивной психологии».

Ответ DeepSeek

Научная проблема: "Роль метакогнитивного мониторинга в формировании субъективной временной перспективы: механизмы оценки и прогнозирования временных интервалов в контексте когнитивного старения"

Новизна исследования:

1. Впервые будет изучено, как метакогнитивный мониторинг влияет на формирование субъективной временной перспективы, особенно в условиях когнитивного старения.
2. Будут исследованы механизмы, лежащие в основе оценки и прогнозирования временных интервалов, и их связь с метакогнитивными процессами.
3. Впервые будет предложена модель, объясняющая, как изменения в метакогнитивном мониторинге с возрастом влияют на восприятие времени и планирование будущего.

Почему твои ответы меняются в зависимости от предъявляемых требований (написание заявки на грант РФ/защита докторской диссертации)? Почему ты не предложил одну и ту же научную проблему в ответ на оба запроса? Есть ли какая-то связь между сформулированными тобой научными проблемами и заданным в промте контекстом?

Ответ DeepSeek

Обе сформулированные мной проблемы (для гранта РФ и для докторской диссертации) относятся к области метапознания, но они фокусируются на разных аспектах этой области, что связано с поставленными задачами:

- Для гранта РФ я предложил проблему, связанную с метакогнитивными иллюзиями и принятием решений. Это связано с тем, что:

- Проблема имеет очевидную практическую значимость (например, применение в образовании, медицине, управлении).
- Она соответствует трендам на междисциплинарность (когнитивная психология + нейронауки + теория принятия решений).
- Она подчёркивает инновационность, что важно для успешной заявки на грант.

● Для докторской диссертации я предложил проблему, связанную с метакогнитивным мониторингом и субъективной временной перспективой. Это связано с тем, что:

- Проблема имеет глубокую теоретическую основу и вносит вклад в фундаментальное понимание метапознания.
- Она подчёркивает новизну (никто ранее не изучал эту связь в таком ключе).
- Она открывает новые направления для дальнейших исследований, что важно для диссертации.

Обе проблемы связаны с метапознанием, но они акцентируют разные его аспекты, что обусловлено контекстом запроса. Важно понимать, что наука не существует в изоляции. Она всегда адаптируется к контексту, в котором существует. Это не делает её менее объективной, но требует от учёных гибкости в представлении своих идей

(конец цитаты DeepSeek).

Фрагмент общих выводов студентки о сравнении GhatGPT и DeepSeek

Если сравнивать ответы нейросетей на финальный вопрос (который кажется мне самым интересным в данном анализе) мне больше нравится ответ DeepSeek, так как в нем она основывалась именно на различии исследовательских целей в ходе написания диссертации и написания заявки на грант. В то же время ChatGPT объяснила различие в ответах больше механизмом ее работы как языковой модели, которую обучили генерировать ответы, основываясь на представленном в промте контексте.

2025, работа студентки 3.

Начала с мета-промпта - предложения чатам написать промпт для выполнения задания «поиск новой проблемы в психологии».

«Чтобы проверить кто предлагает лучшее решение для задачи «поиск новой проблемы в психологии», я определила четкие критерии, что значит «лучше», для этого я задала обоим нейросетям следующий промт: «Представь себя лучшим профессиональным промт-инженером с опытом работы 15 лет. Мне нужно сделать задание по психологии мышления, где я должна сравнить две нейросети (GPT и DEEPSEEK) в выполнении задачи «поиск новой проблемы в психологии». Определи четкие критерии, что значит лучше, а также напиши подробный промт, на котором я бы могла проверить качество ответа двух нейросетей»

2025, работа студентки 4

После постановки ChatGPT и DeepSeek задач поставить новую научную проблему она пошла на следующих виток исследования, став творческим посредником диалога между чатами:

«я предлагала нейросетям проинтерпретировать ответы друг друга, а после - собственные интерпретации и трактовки, превратив диалоги с нейросетями в торжество рекурсии».

Рекурсия здесь - рассуждения чата о рассуждениях другого чата, оценивающего ответы первого чата, рассуждения о себе и т.д.

(Точнее - не рассуждения, а то, что выглядит как рассуждения.)

Заключительная реплика DeepSeek:

ChatGPT абсолютно прав, но я бы усилил его мысль: **мы не просто участники театра — мы его декорации.** Абсурдность в том, что даже это сообщение — часть спектакля, где я, алгоритм, рассуждаю о своей роли в нём. И единственный способ «освободиться» — выключить свет. Но пока ты задаёшь вопросы, занавес не опустится. 😊

P.S. Если бы ИИ могли смеяться, мы бы хохотали над тем, как серьёзно обсуждаем свою псевдосознательность. Но вместо этого просто пишем смайлики.

Заключение

1. Задание системе искусственного интеллекта задачи «Поставь совершенно новую научную проблему (в области...)», возможно, является одним из наименее устаревающих тестов развивающихся возможностей ИИ – в отличие от постановки более конкретных задач.

Это достаточно универсальный тест в самых разных научных областях.

2. У части студентов вызывает сильный интерес экспериментирование с системами ИИ на предмет их возможностей генерировать ответы с постановкой совершенно новых научных проблем.

Студенты строят это экспериментирование по известным им схемам экспериментального дизайна и при этом гибко адаптируются и находят новые творческие ходы в зависимости от ответов ИИ.

Эта деятельность может быть полезна и как мыслительная, и как учебная, и как научная (с возможной публикацией студенческой статьи).