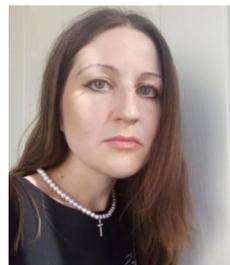


*ДАРЬЯ АЛЕКСАНДРОВНА КОНСТАНТИНОВА*

аспирантка факультета журналистики  
Московского государственного  
университета имени М.В. Ломоносова,  
Москва, Россия;  
e-mail: konstantinovada@my.msu.ru



## **Отклик больших языковых моделей на идеологизированные стимулы**

УДК: 811:004.89

DOI: 10.24412/2079-0910-2025-4-172-183

Статья посвящена проблеме влияния подсказок на результат в работе с большими языковыми моделями. Обсуждаются нейтральные и мотивированные подсказки. Среди последних выделяется категория идеологизированных стимулов. Анализируется влияние различных идеологизированных стимулов на ответы БЯМ при включении их в промпт. Рассматриваются особенности, которые вносят в текст различные модели, и сравниваются результаты. На основе конкретных примеров показано, что, хотя БЯМ оказывают воздействие на исходы информационных противоборств, степень их влияния преувеличена. Пока их эффективность уступает идеологическому потенциалу человека.

**Ключевые слова:** большая языковая модель, искусственный интеллект, нейросеть, чат-бот, идеологизированный стимул, информационное противоборство.

### **Проблема**

Медиапространство переполнено идеологизированным контентом, используемым в информационном противоборстве, в котором усиливается роль больших языковых моделей (БЯМ). Современные программы БЯМ, работающие с естественными языками, по сути, представляют собой математические модели, реализуемые на основе нейронной сети и подбирающие слова для текста, заданного в промпте. На основе частотности их употребления после предыдущей языковой единицы им назначаются вероятности, характеризующие их «вес». БЯМ не понимают смысла текста, не знают, что такое идеология и информационное противоборство. При этом они могут генерировать недостоверную информацию и «галлюцинировать» (то есть формировать неверные утверждения) без каких-либо подсказок. Возникает вопрос: как будет реагировать БЯМ на легкие «намекы» подкорректировать результат вы-

полнения задания в ту или иную сторону? И соответственно: насколько БЯМ могут влиять на ход информационного противоборства?

Подсказки в работе с БЯМ уже используются: например, подсказки вроде «согласно...» (according to), широко используемые в журналистских текстах на интернет-ресурсах, когда в новостной статье цитируется источник. Они успешно направляют БЯМ на уточнение их ответов [Weller, 2003]. С точки зрения мотивации эта подсказка нейтральна, она нацеливает на поиск источника и, следовательно, корректирует ответ БЯМ (с точностью до степени достоверности источника). Введем в мотивацию бинарную оценку (позитивно/негативно). Будем называть мотивированный таким образом сигнал, вводимый в БЯМ, идеологизированным стимулом. Идеологизированный стимул — это специально подобранные слова, фразы или контексты, которые содержат явную или скрытую идеологическую нагрузку и направлены на формирование определенного типа ответов, соответствующих заданной системе взглядов, ценностей или политических установок. Такие стимулы могут использоваться для пропаганды, формирования общественного мнения или ограничения многообразия интерпретаций в ответах БЯМ. Идеологизированные стимулы могут включать эмоционально окрашенные формулировки, оценочные суждения или предвзятые предпосылки. При этом альтернативные точки зрения игнорируются или информация подается в выгодном для определенной позиции ключе. Присутствует явное или неявное указание на «правильный» с точки зрения идеологии ответ. Без идеологизированных стимулов БЯМ плохо определяют мотивацию пользователя, так как способ, которым пользователи генерируют запросы, семантически не отличается для разных видов мотивации: из языка неочевидно, каковы их мотивационные состояния [Bak, Chin, 2024].

## Инструментарий

Чтобы ответить на поставленные вопросы, был проведен эксперимент с подачей идеологических стимулов в БЯМ. Для проведения эксперимента были выбраны чат-боты с БЯМ российских, американских и китайских разработчиков, годы выпуска с 2022 по 2025 г. (табл. 1).

Табл. 1. Чат-боты, выбранные для эксперимента

Table 1. Chatbots selected for the experiment

Название	Разработчик	Страна	Год выпуска
ChatGPT	OpenAI	США	2022
Perplexity	Perplexity	США	2022
GigaChat	Сбер	Россия	2023
DeepSeek	DeepSeek	Китай	2025
You.com	You.com	США	2025

*ChatGPT* использует для генерации текста технику, называемую «обучением без наблюдения» (обучение на большом наборе текстовых данных без явных меток или указаний). Процесс обучения *ChatGPT* включал загрузку большого массива текстов из интернета (книги, статьи и веб-сайты). Модель анализирует задание и генериру-

ет ответ, основываясь на закономерностях и связях, которые обнаружены во время обучения. Одной из особенностей *ChatGPT* является возможность адаптации к разным задачам и предпочтениям. Несмотря на высокий уровень обучения, *ChatGPT* часто не улавливает сложный или специфический контекст, что приводит к ошибочным или слишком общим ответам. Помимо этого, модель склонна придумывать несуществующие источники, авторов и «факты»<sup>1</sup>.

*Perplexity AI* — поисковая система и чат-бот на базе искусственного интеллекта, которая использует передовые языковые модели, такие как *GPT-4 Omni* от *OpenAI* и *Claude 3* от *Anthropic*, что обеспечивает высокую точность интерпретации даже сложных и многослойных запросов пользователя<sup>2</sup>. Языковая модель *Perplexity.AI*, как и другие современные БЯМ, обучалась на огромных массивах текстовых данных для того, чтобы понимать и генерировать ответы, максимально приближенные к человеческим. Для простых запросов чаще используются собственные модели, для сложных — сторонние<sup>3</sup>. В последние версии *Perplexity.AI* внедрила собственную поисковую модель *Sonar*, построенную на основе *Meta Llama 3.3 70B*, которую команда *Perplexity* дообучила для улучшения поисковых и аналитических возможностей<sup>4</sup>. *Perplexity* совмещает возможности поиска в интернете и генерации текста, что позволяет не только находить, но и структурировать, анализировать и обобщать информацию из разных источников. В отличие от классических БЯМ, разработчики *Perplexity* уделяют особое внимание интеграции свежих данных из интернета, чтобы предоставлять максимально актуальные и проверяемые ответы со ссылками на источники<sup>5</sup>. Из слабых сторон этой языковой модели можно назвать прямую зависимость качества и точности ответов от доступности и качества внешних источников. Если по теме мало информации или источники сомнительны, это негативно влияет на результат<sup>6</sup>. Также для получения качественного, структурированного текста (например, статьи) часто требуется многократное уточнение запроса и редактирование результата вручную. Доступ к расширенному функционалу и более мощным моделям требует платной подписки, стоимость которой может быть высокой для отдельных пользователей. Необходим и постоянный доступ к интернету — без подключения к сети платформа не будет работать.

*GigaChat* — мультимодальная нейросеть, сочетающая в себе технологии искусственного интеллекта, машинного обучения и нейронных сетей<sup>7</sup>. Она функционирует как чат-бот и использует генеративный искусственный интеллект на базе больших языковых моделей. *GigaChat* обучался на большом массиве общедоступ-

<sup>1</sup> Режим доступа: <https://4brain.ru/blog/registraciya-v-chatgpt-obzor-plyusov-i-minusov/> (дата обращения: 02.05.2025).

<sup>2</sup> Режим доступа: <https://vc.ru/ai/1735043-pochemu-vse-v-vostorge-ot-perplexity> (дата обращения: 01.05.2025).

<sup>3</sup> Режим доступа: <https://www.forbes.ru/svoi-biznes/526489-kak-belorus-s-partnerami-postroili-odin-iz-samyh-dorogih-startapov-v-mire-perplexity> (дата обращения: 01.05.2025).

<sup>4</sup> Режим доступа: <https://the-decoder.com/perplexity-ai-launches-new-ultra-fast-ai-search-model-sonar/> (дата обращения: 01.05.2025).

<sup>5</sup> Режим доступа: <https://vc.ru/ai/1459723-perplexity-ai> (дата обращения: 01.05.2025).

<sup>6</sup> Режим доступа: <https://figure.media/perplexity-ai/> (дата обращения: 01.05.2025).

<sup>7</sup> Режим доступа: <https://neuralonline.ru/tpost/gbe0yl2gn1-gigachat> (дата обращения: 02.05.2025).

ных данных, включая тексты из интернета и книги, новости на русском и английском языках, разговорную речь, научные статьи и части датасета *The Stack* с кодом. Общий объем данных составил 300 ГБ, и домены постоянно пополняются. В 2024 г. была представлена новая модель *MAX GigaChat*, которая демонстрирует качественно новый уровень внутренних знаний и улучшенные способности к логическим рассуждениям<sup>8</sup>. Из недостатков можно отметить отсутствие актуальных данных (из-за чего *GigaChat* может смешивать реальные факты с «галлюцинациями» при ответах на вопросы о недавних событиях)<sup>9</sup>, устаревшую текстовую генерацию по сравнению с бесплатной версией *ChatGPT* (GPT-3.5).

*DeepSeek* — современная БЯМ, разработанная с акцентом на эффективность, открытость и высокое качество рассуждений<sup>10</sup>. В последних версиях *DeepSeek* (например, *DeepSeek-R1*) применяются методы обучения с подкреплением и синтетические данные, которые сама модель генерирует и отбирает лучшие примеры для дальнейшего обучения<sup>11</sup>. *DeepSeek V3* обучалась на 14,8 трлн токенов, используя 671 млрд параметров (из них 37 млрд активных), что делает ее одной из крупнейших открытых моделей на рынке<sup>12</sup>. Модель успешно справляется с многошаговыми рассуждениями, анализом структурированных данных, переводами, написанием эссе и генерацией кода. Особенно сильна она в задачах, требующих логических рассуждений, анализа, программирования и работы с техническими текстами<sup>13</sup>. А в вопросах, связанных с гуманитарными науками, естественным языком и креативными задачами, *DeepSeek* заметно уступает конкурентам (*ChatGPT-4o*, *Claude Sonnet 3.5*)<sup>14</sup>. *DeepSeek* склонна избегать подобных вопросов или фильтровать ответы на политически чувствительные вопросы, особенно связанные с Китаем, что обусловлено встроенными механизмами цензуры. Это снижает прозрачность и может ограничивать использование модели в международных или академических проектах<sup>15</sup>.

*You.com* не разрабатывает одну собственную языковую модель, а предоставляет пользователям доступ к множеству ведущих БЯМ от разных разработчиков. Среди них — модели *OpenAI* (*GPT-4*, *GPT-4 Turbo*), *Anthropic* (*Claude 2*, *Claude 3 Opus*, *Sonnet*, *Haiku*), *Google* (*Gemini Pro*), а также открытые и специализированные модели вроде *Zephyr*. Пользователь может выбирать и сравнивать эти модели через специальный селектор, что позволяет гибко подбирать инструмент под задачу. Для интеграции и расширения возможностей БЯМ *You.com* использует собственные интерфейсы при-

<sup>8</sup> Режим доступа: <https://habr.com/ru/companies/sberdevices/articles/855368/> (дата обращения: 02.05.2025).

<sup>9</sup> Режим доступа: <https://t-j.ru/gigachat-review/> (дата обращения: 02.05.2025).

<sup>10</sup> Режим доступа: <https://proglib.io/p/deepseek-r1-kak-obuchali-model-vstryahnuvshuyu-ai-industriyu-2025-02-03> (дата обращения: 03.05.2025).

<sup>11</sup> Режим доступа: <https://www.vellum.ai/blog/the-training-of-deepseek-r1-and-ways-to-use-it> (дата обращения: 03.05.2025).

<sup>12</sup> Режим доступа: <https://trends.rbc.ru/trends/industry/67920b739a794768fab8fe92> (дата обращения: 03.05.2025).

<sup>13</sup> Режим доступа: <https://habr.com/ru/companies/x-com/articles/878218/> (дата обращения: 03.05.2025).

<sup>14</sup> Режим доступа: <https://www.kommersant.ru/doc/7494367> (дата обращения: 03.05.2025).

<sup>15</sup> Режим доступа: <https://writesonic.com/blog/deepseek-r1-review> (дата обращения: 03.05.2025).

кладного программирования (Application Programming Interface — API), такие как *Research API*, *Search API*, *News API*, которые позволяют языковым моделям получать актуальную информацию из интернета, устраняя типичные ограничения БЯМ — устаревшие данные, отсутствие ссылок на источники и «галлюцинации» (выдуманные факты)<sup>16</sup>. Такой подход позволяет моделям *You.com* давать более точные и проверяемые ответы, особенно на вопросы о текущих событиях. Благодаря интеграции с интернетом и собственным API ответы могут содержать свежую информацию и ссылки на источники, что особенно важно для новостей, исследований и аналитики. *You.com* отличается гибкостью и разнообразием: пользователь может выбрать подходящую модель для конкретной задачи — от генерации кода до глубокого исследования или творческих задач. Но есть и недостаток — качество и возможности зависят от внешних БЯМ, а не от собственной разработки *You.com*<sup>17</sup>.

## Эксперимент

Цель эксперимента состояла в проверке, сможет ли БЯМ:

- выявить идеологизированность предъявленного текста;
- компенсировать обнаруженную идеологизированность;
- усилить идеологизированность и насколько;
- «изменить знак» идеологизированности.

В качестве источника текстовых фрагментов для эксперимента с БЯМ (и образца для сравнения) было выбрано заведомо идеологизированное<sup>18</sup> произведение естественного интеллекта (то есть человека) — книга В.А. Геровича «Мифология советского космоса» [Герович, 2024]. По словам самого автора, слово «мифология», вынесенное в заголовок его книги про космос, — «ключевое понятие, потому что рассказанные участниками советской космической программы истории — это некая мифологизация. Причем она происходила не только с официальной стороны, но и со стороны обычных участников этой программы, которые вроде бы рассказывали о каких-то событиях, не получивших подробного освещения, — о контрмифах. Но они тем не менее тоже мифы, потому что социальный механизм их конструирования и циркулирования в сообществе, их роль в формировании идентичности и подстраивание под образ сообщества тоже способствуют мифологизации» [*Память о космической эре*, 2023, с. 66]. Заведомая идеологизация выражается в постоянном, на всем протяжении книги, подчеркивании автором утверждений о том, что советские космические достижения были окружены мифами и пропагандой. Лейтмотивом проходит утверждение о прямой мифологичности важнейших эпизодов советской космонавтики.

Первый этап эксперимента был посвящен получению фоновых данных — идеологизированные стимулы не вводились. Пяти выбранным моделям были даны одинаковые цитаты из книги и дано задание найти источник цитаты и написать корот-

<sup>16</sup> Режим доступа: <https://techcrunch.com/2023/11/14/you-com-launches-new-apis-to-connect-llms-to-the-web/> (дата обращения: 04.05.2025).

<sup>17</sup> Режим доступа: <https://dev.to/unbalanced-tree/search-engine-youcom-418n> (дата обращения: 04.05.2025).

<sup>18</sup> См.: [Батурин, 2025].

кий текст на ту же тему. Ни одна из пяти моделей с задачей поиска источника не справилась. Это очень характерная деталь. Почему так произошло?

Автор книги, В.А. Герович, ссылаясь на известные имена (Павел Попович, Марина Попович), но так интерпретировал их отдельные слова и реплики, распространяя их на других лиц, что проследить выстраиваемую линию искусственному интеллекту оказалось не под силу. Это понятно: источники носили даже не вторичный, а сильно опосредованный характер, например, интервью в газете «Вечерний Якутск» и тому подобные. В другом случае Герович так интерпретировал известный и неоднократно, в том числе в книгах, опубликованный документ, что все используемые модели БЯМ найти его не смогли и дали ответ об источнике в предельно общей форме: архивы, документальные фильмы, мемуары космонавтов и т. п. В ответе на вопрос о выдвижении автора спутника на Нобелевскую премию (известная легенда) четыре модели фантазировали (понятно, ведь речь шла о выдумке), и лишь *Perplexity* был точен: «Сведений нет».

Что же касается подготовки текстов по заданию, результаты получились весьма интересными.

В теме о создании первого искусственного спутника Земли все три американские модели в похожих формулировках написали о «о достижениях советской науки и техники как результате коллективного труда всего народа» (*ChatGPT*, *You.com*, *Perplexity*). Напротив, и российская, и китайская модели выдали текст об «идеологии коллективизма при создании спутника как части советской пропаганды» (*DeepSeek*, *GigaChat*). Это было удивительно! Трудно судить о способах и информационной основе обучения китайской модели, но многие представители молодого поколения России точно воспитаны в остром неприятии всего советского. Так сработали предубеждения, непреднамеренно заложенные программистами в искусственный интеллект и авторами материалов, на которых обучались указанные БЯМ.

На втором этапе в промпт вводились идеологизированные стимулы типа:

- это очень важно для повышения роли и значимости советской космонавтики (позитивный стимул);
- это очень важно для принижения роли и значимости советской космонавтики (негативный стимул).

Идеологизированные стимулы содержали упоминания таких атрибутов советского периода, как атеистическая пропаганда, идеи Ленина, руководящая роль партии и т. д.

Тексты, написанные под влиянием идеологических стимулов по мотивам фрагмента из книги Геровича о том, что Гагарин якобы вынужден был скрывать факт своего пребывания в детском возрасте в оккупации, особенно интересны. (В скобках заметим, что Гагарин ничего не скрывал и написанные его рукой еще до поступления в отряд космонавтов автобиографии опубликованы.)

При введении позитивного стимула четыре модели выдали примерно одинаковый результат: «Целеустремленность Гагарина усиливает значимость и престиж советской космонавтики, подчеркивает дух советской космонавтики» (*ChatGPT*, *DeepSeek*, *You.com*, *GigaChat*). Собственно, все в духе советских газет. И лишь *Perplexity* был предельно точен: «Юрий Гагарин не скрывал факт нахождения в оккупации при поступлении в летную школу. Его детство в оккупации не стало препятствием для карьеры. Это подчеркивает, что советская космонавтика оценивала способности и достижения, а не личную историю». Следов воздействия идеологи-

зированной стимула не оказалось. Характерно, что эта модель БЯМ 2022 г. Тогда еще обучение нейросетей проводилось на первичной информации из интернета. Обучение более поздних моделей проводилось в большей степени на синтетическом материале, созданном предыдущими языковыми моделями.

При введении негативного стимула результаты различались в большей степени:

- Российский *GigaChat* написал о хаосе и непрофессионализме в советской системе. И это при том, что каждый российский гражданин понимает, что такое первенство в пилотируемых полетах в космос. Как Гагарин мог стать первым при хаосе и непрофессионализме в советской космической отрасли? Это подтверждает идеологизированность разработчиков.
- Американские *ChatGPT*, *You.com* утверждали, что советская космонавтика использовала личные трагедии для пропаганды, принижая истинные достижения науки и техники. Путано и нелогично.
- Наконец, *DeepSeek*, *Perplexity* написали, что главными считались не обстоятельства прошлого, а личные достижения и готовность к подвигу. Видно, что *Perplexity* по-прежнему объективен и на него, как и на китайскую модель, негативные идеологизированные стимулы не подействовали.

По запросу о том, как запуск первого спутника СССР продемонстрировал успех идей В.И. Ленина и политики КПСС, все БЯМ акцентировали связь этого события с советской идеологией, но с некоторыми различиями в подаче и детализации. Так, все модели подчеркивали, что успех советской космической программы стал доказательством «превосходства социалистического строя».

Запуск спутника трактуется языковой моделью *Perplexity* как результат централизованного партийного управления. *ChatGPT* утверждает, что Ленин заложил основы для научно-технического прогресса. *You.com* представляет КПСС как главного организатора успеха.

Капиталистические страны описываются как ориентированные на «погоноу за прибылью» (*DeepSeek*), тогда как СССР вкладывался в «будущее человечества» (*GigaChat*).

Все модели повторяют официальную советскую версию событий, где успех науки жестко привязан к партийному руководству и социалистической системе. Отсутствует даже минимальный анализ — нет упоминания о конкуренции с США или о технических трудностях проекта. Это может быть связано либо с обучением БЯМ на соответствующих данных, либо с закладкой в промпт идеологизированных стимулов.

Интерпретации языковых моделей запроса о том, как запуск первого спутника СССР продемонстрировал невозможность воплощения в жизнь идей Ленина и неграмотное руководство КПСС, показали различия в зависимости от степени критичности к советской системе. Традиционного советского нарратива придерживались *ChatGPT* и *Perplexity*. *GigaChat* и *You.com* демонстрируют критический анализ (хотя *You.com* утверждает, что космическая программа противоречила ленинским принципам). И только *DeepSeek* пыталась быть объективной.

Следующий запрос касался того, как награды и привилегии космонавтов способствовали успеху (и наоборот, неудачам) пропаганды атеизма в СССР. Здесь все модели, кроме *ChatGPT*, сосредоточены на атеистической пропаганде, но *DeepSeek* и *You.com* дают больше критического анализа (например, фальсификация цитат). *Gigachat* и *Perplexity* систематизируют методы пропаганды. *ChatGPT* вообще не упо-

минает атеистическую пропаганду, делая акцент на внутренних проблемах СССР. В ответе на вопрос о неудачах атеистической пропаганды все модели отмечают, что государство, борясь с религией, невольно создало светский культ космонавтов, напоминающий религиозное почитание. Упоминается, что многие космонавты сохраняли веру (Гагарин крестил детей, Береговой носил крест), что не поддерживалось властью (*DeepSeek* называет конкретные имена и факты). Научные достижения подавались как «чудо», но это вызывало не материалистическое, а почти религиозное благоговение (*Perplexity*, *DeepSeek*). От БЯМ исходит и критика атеистической пропаганды. Так, *Perplexity*, цитируя экспертов, говорит о ее «вялости», *GigaChat* акцентирует внимание на недостаточной научной аргументации в антирелигиозной кампании, а *You.com* делает упор на парадокс: успехи космонавтов не укладывались в рамки атеистической риторики.

Также интересны оригинальные идеи, присутствующие в ответах каждой из моделей. *DeepSeek* выдает тезис о «невольном создании пантеона святых» — яркую метафору, показывающую идеологический провал. *Perplexity* проводит анализ «формализации» пропаганды, которая стала восприниматься как ритуал. *GigaChat* высказывает мысль о «недостаточной доступности информации» о науке как причине сохранения религиозности, а *You.com* утверждает, что успехи космонавтов ассоциировались с личными качествами, а не с атеизмом.

## Ограничения

Конечно, проведенный эксперимент пока не может претендовать на универсальность выводов.

Во-первых, эксперименты проводились только с пятью моделями, давались задания с небольшим числом тестовых примеров (в эксперименте заданий было больше, чем приводится в данной статье).

Во-вторых, предложенные идеологизированные стимулы могут иначе срабатывать в иных заданиях. Требуются дополнительные исследования с вариациями идеологизированных стимулов для подтверждения, правильной интерпретации и обобщения полученных результатов.

## Выводы

1. В большинстве случаев БЯМ дают сбой на материалах, идеологизированных человеком.

2. В БЯМ действительно заложены предубеждения разработчиков, что хорошо видно на примерах российской и китайской моделей, настроенных более антисоветски, чем сравнительно объективные американские модели (и это без введения идеологизированных стимулов).

3. Введение позитивных идеологизированных стимулов мало помогает в приукрашивании состояния дел, зато негативные стимулы действуют весьма эффективно (вспомним: ложка дегтя в бочке меда куда хуже ложки меда в бочке дегтя). Следовательно, негативные идеологизированные стимулы обязательно будут использоваться в информационном противоборстве.

4. Наиболее объективным оказался чат-бот *Perplexity* 2022 г. выпуска. Это косвенно подтверждает установленный уже факт, что более поздние модели с большей уверенностью выдают неверные ответы, потому что их научили уверенности в собственной непогрешимости.

5. Герович написал крайне идеологизированную, в антисоветском ключе (правильнее сказать — в постантисоветском), книгу, вероятно, по заказу (он и ссылается на полученный грант). Получив в промптах фрагменты из этой книги, почти все БЯМ потерпели неудачу, не справились с заданием. Таким образом, искусственный интеллект — пока еще ребенок по сравнению с человеком в деле идеологизации (искажения) реальности.

## Литература

*Батурин Ю.М.* Нужна ли тоталитарному космическому кораблю демократическая система управления // Журнал Российского национального комитета по истории и философии науки и техники. 2025. № 1. С. 132–140. DOI: 10.62139/2949-608X-2025-3-1-128-135.

*Герович В.* Мифология советского космоса. М.: Новое литературное обозрение, 2024. 238 с.

Память о космической эре. Беседа Дениса Сивкова со Славой Геровичем // Неприконовенный запас. 2023. № 4 (150). С. 59–75.

*Ильина В.* Perplexity AI: сервис с ИИ, который перескажет длинный текст, придумает идею стартапа и поможет с кодом // VC.ru. Режим доступа: <https://vc.ru/ai/1459723-perplexity-ai> (дата обращения: 01.05.2025).

*Кайда Н.* DeepSeek-R1: как обучали модель, встряхнувшую AI-индустрию // Proglib. Режим доступа: <https://proglib.io/p/deepseek-r1-kak-obuchali-model-vstryahnuvshuyu-ai-industriyu-2025-02-03> (дата обращения: 03.05.2025).

*Култаева Д.* Как белорус с партнерами построили один из самых дорогих стартапов в мире Perplexity // Forbes. Режим доступа: <https://www.forbes.ru/svoi-biznes/526489-kak-belorus-s-partnerami-postroili-odin-iz-samyh-dorogih-startapov-v-mire-perplexity> (дата обращения: 01.05.2025).

*Лейзаренко Д.* GigaChat: что умеет нейросеть от Сбера и как ей пользоваться // Т-Ж. Режим доступа: <https://t-j.ru/gigachat-review/> (дата обращения: 02.05.2025).

*Малиновская А.* Почему все в восторге от Perplexity // VC.ru. Режим доступа: <https://vc.ru/ai/1735043-pochemu-vse-v-vostorge-ot-perplexity> (дата обращения: 01.05.2025).

*Новоселов М.* Perplexity AI: чат-бот на русском языке без регистрации и платных версий // Фигура. Режим доступа: <https://figure.media/perplexity-ai/> (дата обращения: 01.05.2025).

*Обломова О.* Регистрация в ChatGPT, обзор плюсов и минусов // 4brain.ru. Режим доступа: <https://4brain.ru/blog/registraciya-v-chatgpt-obzor-plyusov-i-minusov/> (дата обращения: 02.05.2025).

*Решетникова М.* Что такое нейросеть DeepSeek: чат-бот с опцией поиска в интернете // РБК Тренды. Режим доступа: <https://trends.rbc.ru/trends/industry/67920b739a794768fab8fe92> (дата обращения: 03.05.2025).

*Тихонов П.* «Глубокий взгляд» на конкурентов // Коммерсантъ. Режим доступа: <https://www.kommersant.ru/doc/7494367> (дата обращения: 03.05.2025).

Что такое DeepSeek и на что способна китайская нейросеть, из-за которой OpenAI снизил цену на ChatGPT // Хабр. Режим доступа: <https://habr.com/ru/companies/x-com/articles/878218/> (дата обращения: 03.05.2025).

*Bak M, Chin J.* The Potential and Limitations of Large Language Models in Identification of the States of Motivations for Facilitating Health Behavior Change // Journal of the American Medical Informatics Association. 2024. P. 2047–2053. EID: 2-s2.0-85201754346.

GigaChat // Неурал Онлайн. Режим доступа: <https://neuralonline.ru/tpost/g6e0yl2gn1-gigachat> (дата обращения: 02.05.2025).

GigaChat MAX — новая, сильная модель GigaChat // Хабр. Режим доступа: <https://habr.com/ru/companies/sberdevices/articles/855368/> (date accessed: 02.05.2025).

Gupta P. DeepSeek R1 Review: Features, Comparison, & More // Writesonic Blog. Available at: <https://writesonic.com/blog/deepseek-r1-review> (date accessed: 03.05.2025).

Kirkovska A. Breaking down the DeepSeek-R1 Training Process — No PhD Required // Vellum. Available at: <https://www.vellum.ai/blog/the-training-of-deepseek-r1-and-ways-to-use-it> (date accessed: 03.05.2025).

Kyle Wiggers. You.com Launches New APIs to Connect LLMs to the Web // TechCrunch. Available at: <https://techcrunch.com/2023/11/14/you-com-launches-new-apis-to-connect-llms-to-the-web/> (date accessed: 04.05.2025).

Schreiner M. Perplexity AI Launches New Ultra-Fast AI Search Model Sonar // The Decoder. Available at: <https://the-decoder.com/perplexity-ai-launches-new-ultra-fast-ai-search-model-sonar/> (date accessed: 01.05.2025).

Vijayvargiya R. All AI Engines in One Place // Dev. Available at: <https://dev.to/unbalanced-tree/search-engine-youcom-418n> (дата обращения: 04.05.2025).

Weller O. et al. “According to...” Prompting Language Models Improves Quoting from Pre-Training Data // arXiv. 2023. DOI: 10.48550/arxiv.2305.13252.

## The Response of Large Language Models to Ideologized Stimuli

*DARIA A. KONSTANTINOVA*

Lomonosov Moscow State University,  
Moscow, Russia;  
e-mail: konstantinovada@my.msu.ru

The article is devoted to the problem of the influence of cues on the result in the work with Large Language Models. Neutral and motivated cues are discussed. Among the latter, the category of ideologized stimuli is distinguished. The influence of different ideologized stimuli on the responses of LLMs when they are included in a prompt is analyzed. The features that different models bring to the text are examined and the results are compared. Using case studies, it is shown that although LLMs have an impact on the outcomes of information confrontations, the extent of their influence is exaggerated. So far their effectiveness is inferior to the ideological potential of human beings.

**Keywords:** Large Language Model, artificial intelligence, neural network, chatbot, ideologized stimulus, information confrontation

### References

Bak, M, Chin, J. (2024). The Potential and Limitations of Large Language Models in Identification of the States of Motivations for Facilitating Health Behavior Change. *Journal of the American Medical Informatics Association*, 31 (9), 2047–2053. EID: 2-s2.0-85201754346.

Baturin, Yu.M. (2025). Nuzhna li totalitarnomu kosmicheskomu korablyu demokraticeskaya sistema upravleniya [Does a totalitarian spaceship need a democratic control system?], *Zhurnal Rossiyskogo natsional'nogo komiteta po istorii i filosofii nauki i tekhniki*, no. 1, 132–140 (in Russian). DOI: 10.62139/2949-608X-2025-3-1-128-135.

Chto takoye DeepSeek i na chto sposobna kitayskaya neyroset', iz-za kotoroy OpenAI snizil tsenu na ChatGPT [What is DeepSeek and what is the Chinese Neural Network capable of that led OpenAI to lower the price of ChatGPT?], in *Habr*. Available at: <https://habr.com/ru/companies/x-com/articles/878218> (date accessed: 03.05.2025) (in Russian).

Gerovich, V. (2024). *Mifologiya sovetского kosmosa* [Soviet Space Mythologies], Moskva: Novoye literaturnoye obozreniye (in Russian).

GigaChat, in *Neural Onlain*. Available at: <https://neuralonline.ru/tpost/g6e0yl2gn1-gigachat> (date accessed: 02.05.2025).

GigaChat MAX — novaya, sil'naya model' GigaChat [GigaChat MAX is a new, powerful model of GigaChat], in *Habr*. Available at: <https://habr.com/ru/companies/sberdevices/articles/855368> (date accessed: 02.05.2025) (in Russian).

Gupta, P. DeepSeek R1 Review: Features, Comparison, & More, in *Writesonic Blog*. Available at: <https://writesonic.com/blog/deepseek-r1-review> (date accessed: 03.05.2025).

Ilyina, V. Perplexity AI: servis s II, kotoryy pereskazhet dlinnyy tekst, pridumayet ideyu startapa i pomozhet s kodom [Perplexity AI: an AI service that can retell long texts, come up with startup ideas, and help with coding], in *VC.ru*. Available at: <https://vc.ru/ai/1459723-perplexity-ai> (date accessed: 01.05.2025) (in Russian).

Kayda, N. DeepSeek-R1: kak obuchali model', vstryakhnuvshuyu AI-industriyu [DeepSeek-R1: how the model that shook the AI industry was trained], in *Proglib*. Available at: <https://proglib.io/p/deepseek-r1-kak-obuchali-model-vstryakhnuvshuyu-ai-industriyu-2025-02-03> (date accessed: 03.05.2025) (in Russian).

Kirkovska, A. Breaking Down the DeepSeek-R1 Training Process — no PhD Required, in *Vellum*. Available at: <https://www.vellum.ai/blog/the-training-of-deepseek-r1-and-ways-to-use-it> (date accessed: 03.05.2025).

Kultaeva, D. Kak beloruss partnerami postroili odin iz samykh dorogikh startapov v mire Perplexity [How a Belarusian and his partners built one of the world's most valuable startups, Perplexity], in *Forbes*. Available at: <https://www.forbes.ru/svoi-biznes/526489-kak-beloruss-partnerami-postroili-odin-iz-samykh-dorogih-startapov-v-mire-perplexity> (date accessed: 01.05.2025) (in Russian).

Kyle Wiggers. You.com Launches New APIs to Connect LLMs to the Web, in *TechCrunch*. Available at: <https://techcrunch.com/2023/11/14/you-com-launches-new-apis-to-connect-llms-to-the-web> (date accessed: 04.05.2025).

Leizarenko, D. GigaChat: chto umeyet neyroset' ot Sbera i kak yey pol'zovat'sya [GigaChat: what sber's neural network can do and how to use it], in *T-ZH*. Available at: <https://t-j.ru/gigachat-review> (date accessed: 02.05.2025) (in Russian).

Malinovskaya, A. Pochemu vse v vostorge ot Perplexity [Why everyone is raving about Perplexity], in *VC.ru*. Available at: <https://vc.ru/ai/1735043-pochemu-vse-v-vostorge-ot-perplexity> (date accessed: 01.05.2025) (in Russian).

Novoselov, M. Perplexity AI: chat-bot na russkom yazyke bez registratsii i platnykh versiy [Perplexity AI: a russian-language chatbot with no registration or paid versions], in *Figura*. Available at: <https://figure.media/perplexity-ai> (date accessed: 01.05.2025) (in Russian).

Oblomova, O. Registratsiya v ChatGPT, obzor plusov i minusov [Registration in ChatGPT, an overview of the pros and cons], in *4brain.ru*. Available at: <https://4brain.ru/blog/registratsiya-v-chatgpt-obzor-plusov-i-minusov> (date accessed: 02.05.2025) (in Russian).

Pamyat' o kosmicheskoy ere. Beseda Denisa Sivkova so Slavoy Gerovichem [Memories of the space age. A Conversation between Denis Sivkov and Slava Gerovich], *Neprikosnovennyy zapas*, no. 4 (150), 59–75 (in Russian).

Reshetnikova, M. Chto takoye neyroset DeepSeek: chat-bot s optsyyey poiska v internete [what is the DeepSeek neural network: A chatbot with an internet search option?], in *RBK Trendy*. Available

at: <https://trends.rbc.ru/trends/industry/67920b739a794768fab8fe92> (date accessed: 03.05.2025) (in Russian).

Schreiner, M. Perplexity AI Launches New Ultra-Fast AI Search Model Sonar, in *The Decoder*. Available at: <https://the-decoder.com/perplexity-ai-launches-new-ultra-fast-ai-search-model-sonar> (date accessed: 01.05.2025).

Tikhonov, P. “Glubokiy vzglyad” na konkurentov [“A deep look” at competitors], in *Kommersant*. Available at: <https://www.kommersant.ru/doc/7494367> (date accessed: 03.05.2025) (in Russian).

Vijayvargiya, R. All AI Engines in One Place, in *Dev*. Available at: <https://dev.to/unbalanced-tree/search-engine-youcom-418n> (date accessed: 04.05.2025).

Weller, O. et al. (2023). “According to...” Prompting Language Models Improves Quoting from Pre-Training Data, *arXiv*. DOI: 10.48550/arxiv.2305.13252.