

БЛОГИ

"Своими" словами. Светлана Прокопьева - об авторстве в эпоху ИИ

Светлана Прокопьева

26 октября 2026



OpenAI внедряет textGrain – невидимые "водяные знаки" в тексты, созданные ChatGPT. Массовая предубежденность естественного интеллекта в отношении искусственного получает твердую почву под ногами: теперь дискриминировать произведения, созданные с помощью нейросети, можно будет с опорой на данные самой нейросети. Но есть нюанс. Маркировка "ИИ-контент" почти ничего не говорит о справедливой доле интеллектуального вклада человека и машины в рамках одного произведения.

Как это работает? ChatGPT, генерируя текст, выбирает каждое следующее слово в определенной статистической вероятности, а textGrain чуть изменяет эту вероятность. Потом этот "водяной знак" может быть считан специальным детектором.

Маркировка внедряется из-за требований европейского законодательства. Компания при этом подчеркивает, что детектор определяет признаки участия технологии, но не устанавливает авторство. То есть текст, который человек написал сам, а потом отдал машине на редактуру, может быть отмечен как ИИ-контент. И наоборот: узнаваемость полностью сгенерированного текста после замены 25% слов падает с 92% до 17%.

Пока что доступ к детектору OpenAI выдает только отдельным доверенным организациям. Но я не удивлюсь, если интернет развернет общественную кампанию за открытое для всех приложение. Тогда текстовые публикации в соцсетях ждет та же судьба, что ИИ-картинки и видеоролики: пессимизация и пренебрежение уже по факту маркировки.

Люди с заметным предубеждением относятся ко всему, что создается с помощью ИИ. Нейрослоп – новое ругательство. Пометка "ИИ-контент" фактически приравнивается к сигналу о недостоверности.

Это не только наблюдение – исследования подтверждают. В 2024 году исследователи из Цюриха провели эксперименты, которые показали, что люди склонны меньше доверять даже правдивым новостным заголовкам, когда считают их созданными ИИ. Исследование 2026 года показало, что публикации в Инстаграме (в эксперименте – имитация таких публикаций) вызывают меньший эмоциональный отклик и вовлеченность, если стоит ИИ-маркировка, в том числе на контенте, полностью созданном человеком.

Истоки предубеждения можно понять. Человечество впервые оказалось в ситуации, когда рядом с ним на планете существует еще один тип интеллекта. Нечто/некто, способный размышлять, делать выводы, решать задачи, исследовать белки и рисовать мультики не хуже, а местами и лучше людей. Монополия человека на статус разумного существа ощутимо пошатнулась. Неприятная конкуренция.



СМОТРИ ТАКЖЕ

Цена за выбор быть собой. Светлана Прокопьева – о деле Льва Шлосберга

Ревность мобилизует поиск уязвимостей, и довольно скоро люди нащупали рациональное основание своей подозрительности: искусственный интеллект может отучить человека думать. Тут тоже есть доказательная база: несколько экспериментов уже показали, что творческие способности студентов падают после использования нейросетей. Школьники, которым доступны чат-боты, получают более низкие оценки при проверке знаний.

Если постоянно делегировать машине формулирование мыслей, действительно, легко утратить этот навык. Но вот следующий переход – использование ИИ равно отказу от самостоятельного мышления – вообще-то не доказан. И, как я думаю, не особо-то обоснован.

Приведу пару примеров, как работает ИИ-дискриминация и почему она не имеет отношения к авторской субъектности.

Первый пример. Реальный случай: человек решает сделать политическое заявление. Он выходит в пикет с плакатом, но условия с обстановкой и освещением не идеальные, фотография получилась неудачной. Человек просит ChatGPT улучшить фото – ярче и заметнее прописать его собственный лозунг на плакате.

Facebook обнаруживает ИИ-обработку и маркирует изображение. Человек в панике удаляет публикацию и заменяет на старую, натуральную, плохо читаемую фотографию.

Стало ли его высказывание достовернее? Технически – да. Стало ли оно эффективнее? Вряд ли. Плохая фотография привлечет меньше внимания, меньше людей прочтут пост, будет меньше реакций – ниже шанс достучаться до адресата.

“ Если человек подписывается под заявлением и говорит: да, я так считаю – какая разница, чем он нанес надпись на свой плакат?

Документальная достоверность фотографии не равна подлинности политического высказывания. Мы уже знаем примеры полностью цифровых пикетов – и совершенно реальных административных разбирательств по их поводу. Информационные кампании сегодня зачастую вообще не имеют выхода в физическую реальность – но добиваются

результатов.

Если человек готов поставить фото на личную страницу в соцсети, подписывается под заявлением и говорит: да, я так считаю – какая разница, маркером, фотошопом

или нейросетью он нанес надпись на свой плакат?

Детектор может обнаружить вмешательство ИИ в изображение. Но принадлежит ли человеку выраженная посредством него политическая позиция, он никак не установит.

Другой пример. Человек пишет пост о том, что человеческие тексты всегда лучше машинных (не только по качеству, но и в плане когнитивной пользы для создателя), и что сам он всегда безошибочно определяет, написал ли текст живой человек или нейросеть. При этом пост демонстрирует все признаки нейрослопа: многословие, повторы, предсказуемую композицию, банальности. Не только я, но и мой ChatGPT уверенно опознал в авторе своего собрата.

Однако автор в ответ на подкол уверяет, что нет, никакого ИИ – все сам, своими ручками. И ведь не обязательно врет! Возможно, мне просто попался тот самый классический пример человеческой графомании, на которой весь ИИ как раз и учился писать. Потому что коллективное человечество в статистически усредненном варианте вообще-то совсем не Маркес и не Довлатов.

Как плохое качество текста не доказывает его машинного происхождения, так и машинное происхождение не обещает плохого качества. Но какая же ирония, что предубеждение против ИИ теперь заставляет живых людей отстаивать происхождение своего словоблудия!

Так что же такое авторство в эпоху ИИ?

Для меня это проявление субъектности. А субъектность в данном случае – способность распоряжаться собственным высказыванием. Во-первых, понимать, что ты вообще сказал. Во-вторых, уметь обосновать и отстоять свою позицию в споре.



СМОТРИ ТАКЖЕ

ИИ говорит "нет". Светлана Прокопьева – о страхе перед сверхразумом

Причем имеет смысл разделять юридическую и философскую сторону авторства.

И тут, если мы совершим небольшую смысловую подмену – подставим вместо слова "искусственный" слово "другой", – то внезапно обнаружим, что проблема участия двух и более интеллектов в создании одного "авторского" произведения стара как мир.

Известный писатель выпускает роман за романом, пользуясь трудом "литературных негров". Именитые живописцы поручают ученикам второстепенные детали своих шедевров. Профессора защищают диссертации, пользуясь трудом ассистентов. Политики вообще читают речи, написанные спичрайтерами. Думаю, сегодня эти спичрайтеры – уже ИИ или как минимум его используют.

С политиками особенно наглядно: они же не просто подписываются под словами с бумажки самым фактом публичного озвучивания. Они берут на себя обязательство следовать своим обещаниям, защищать свои принципы, доказывать свои заявления. Вопрос тут только один: насколько они "свои"?

Насколько ты проживаешь внутри себя мысль, которую выдаешь от своего имени?

“ Автор – не обязательно тот, кто придумал и пальцами набрал каждое слово. Автор – тот, кто сам решает, что именно хочет сказать. И готов отвечать за это

Важен не сам факт использования ИИ в производстве текста или картинки. Важен режим использования ИИ.

Ты можешь попросить ChatGPT написать пост или комментарий на тему и опубликовать его в соцсети, даже не перечитывая. Это режим делегирования мышления. Или ты можешь сказать чату: поспорь со мной. Разверни мою гипотезу, раскритикуй, скажи, в чем я не прав? Это

режим тренажера. И в таком случае тебе придется проделать даже больше интеллектуального труда, чем если бы ты в одиночку напечатал свои размышления.

И в этом смысле автор – не обязательно тот, кто придумал и пальцами набрал каждое слово. Автор – тот, кто сам решает, что именно хочет сказать. И готов отвечать за это.

Водяные знаки нейросетей вообще не отвечают на такие вопросы.

И все-таки, какой ироничный поворот: в страхе потерять свою интеллектуальную самостоятельность люди делегируют технологии суждение о чужой

интеллектуальной самостоятельности. Мы разрешаем машине решать, кто подлинный автор, какое творчество настоящее, какой текст имеет ценность.

Как будто сами уже утратили способность к оценке.

Светлана Прокопьева – журналист Радио Свобода

Высказанные в рубрике "Блоги" мнения могут не отражать точку зрения редакции



Светлана Прокопьева

Псковский журналист, редактор и автор регионального сайта Радио Свобода
Север.Реалии

Блоги

Мнения