

Дискуссия ВОИС
ИС и передовые
технологии

Генеративный ИИ



WIPO

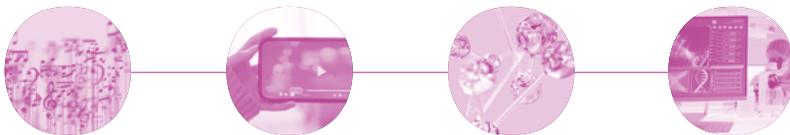
Генеративный ИИ

«Время от времени появляются новые технологии, которые захватывают воображение людей во всем мире, доминируя среди главных новостей и в частных разговорах, разделяя людей на их противников и сторонников и, казалось бы, ставя наш мир на перепутье». Генеральный директор ВОИС Дарен Танг

Что такое генеративный ИИ?

Традиционные системы искусственного интеллекта (ИИ) используются в основном для анализа данных и составления прогнозов. Однако генеративный ИИ расширяет эти возможности, создавая новые данные, аналогичные тем, которые используются для его обучения. Это усовершенствование позволяет создавать различные формы нового контента, включая аудиоматериалы, алгоритмы, изображения, текст, имитационные модели и видеоматериалы.

Генеративный ИИ – это любая модель машинного обучения, способная динамически создавать выходные данные, после того как она соответствующим образом обучена. Важно отметить, что область применения генеративного ИИ выходит далеко за рамки создания контента. Потенциальные области применения генеративных моделей ИИ огромны – от создания музыки и видео до молекулярного моделирования для изыскания новых лекарственных веществ и медицинской диагностики. По мере того как эта технология продолжает развиваться, возможности для ее применения появляются также в других областях.



 В то время как современные модели генеративного ИИ особенно хорошо подходят для сопоставления языков и создания контента, существуют потенциальные возможности для широкого применения генеративного ИИ как в сфере творчества, так и в сфере инноваций.

В результате запуска ChatGPT в ноябре 2022 года генеративный ИИ вышел на передний план, привлек внимание многих и изменил то, как мы работаем и творим. ChatGPT использует обработку естественного языка и глубокое машинное обучение для создания текстового контента. Это была первая в своем роде программа, но вскоре за ней последовали аналогичные модели, в том числе Google Bard, Microsoft Bing Chat и Meta Llama 2. Другие модели искусственного интеллекта, в том числе Mid journey, DALL-E и Stable Diffusion, генерируют изображения с помощью текстовых подсказок.



ChatGPT, первая в своем роде программа, использует обработку естественного языка и глубокое машинное обучение для создания текстового контента.

Хотя некоторые восхищаются точностью и скоростью, с которой генеративный ИИ способен создавать контент, критики обеспокоены потенциальной ненадежностью и ошибками в генерируемых выходных данных. Генеративные модели ИИ базового уровня, используя определенный контекст, такой как фрагмент текста, просто предсказывают наиболее вероятное следующее слово.



Какими бы убедительными ни казались результаты, современные генеративные модели ИИ все-таки имеют серьезные недостатки, поскольку они работают, комбинируя слова в статистически вероятном порядке без подлинного понимания.



Определения

Глубокое обучение, разновидность машинного обучения, опирается на структуру и функционирование человеческого мозга, используя данные для генерирования шаблонов и представлений. При глубоком обучении используются модели нейронной сети с взаимосвязанными узлами, или нейронами, расположенными на нескольких уровнях. Сеть обрабатывает входные данные на этих уровнях и выдает нелинейный результат. У каждого нейрона есть несколько адаптируемых параметров, которые не запрограммированы заранее, а учатся на данных.

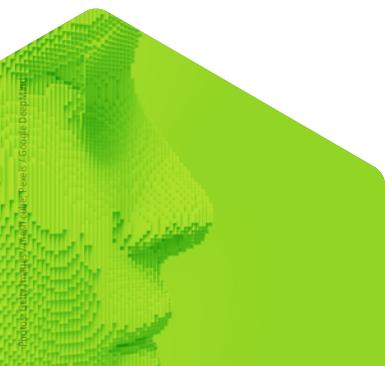
Последние достижения в области генеративного ИИ, которые обычно называют языковыми моделями (LM) или **большими языковыми моделями** (LLM), стали возможны благодаря появлению новых моделей, оптимально подходящих для сопоставления языков. Важно отметить, что LM и LLM не ограничиваются обработкой и генерированием естественных языков. С точки зрения машинного обучения языком могут считаться любые символы (например, слова), которые в определенном контексте (грамматика, отношения между словами) могут передавать значение. В языке слова создают последовательности, и значение зависит не только от выбора слов, но и от отношений между этими словами. Способные эффективно фиксировать эти отношения, LLM подходят для обработки последовательностей и сохранения достаточного объема памяти и предназначены для работы с крупными массивами данных. LLM могут выполнять различные задачи по обработке естественного языка, включая генерацию текста, перевод на другой язык, реферирование текста, ответы на вопросы и многое другое.

Регулирование генеративного ИИ

Генеративный ИИ открывает значительные возможности, но в то же время связан со значительными рисками. Что касается рисков, то он вызывает множество проблем, помимо его технических недостатков, касающихся надежности, точности и этических соображений. Во многих случаях в центре внимания общественных дискуссий оказываются этические проблемы, такие как возможные социальные риски, связанные с дезинформацией, манипулированием рынком, киберпреступлениями, угрозами для неприкосновенности частной жизни и демократии, а также непреднамеренные последствия для создания контента и культурного разнообразия. Кроме того, существуют серьезные опасения по поводу ликвидации рабочих мест, смещения в сторону определенных языков, отсутствия прозрачности, влияния крупных компаний на разработку правил и нормативов, экономического неравенства и монетизации данных, собираемых в странах, не получающих от этого соразмерных выгод.

Многие обеспокоены деструктивным потенциалом генеративного ИИ, в связи с чем необходимы меры регулирования. Они выступают за международное сотрудничество и глобальное управление ИИ. Потребность в быстрых, краткосрочных решениях заставляет политиков рассматривать возможность принятия временных корректирующих мер законодательного характера и искать альтернативные варианты решения потенциальных проблем.

 Генеративный ИИ ставит перед регулирующими органами множество проблем, особенно этических. Регулирующие органы активно пытаются найти быстрые, краткосрочные решения, соответствующие быстрым темпам технического прогресса.



Основные соображения с точки зрения ИС

Краткий обзор

Многие все более широко используют инструменты генеративного ИИ для записей, создания визуального контента и кодирования. Благодаря своей предполагаемой способности самостоятельно создавать контент, генеративный ИИ потенциально способен экономически конкурировать с создателями-людьми, стимулировать творческую деятельность и повышать производительность людей в различных отраслях. Одним из важных примеров, иллюстрирующих потенциальные революционные преобразования в творческих отраслях, является песня «Heart on My Sleeve» («Душа нараспашку»), в которой без разрешения был использован сгенерированный ИИ вокал, имитирующий голоса музыкантов Дрейка и Уикнда (The Weeknd). Это заостряет внимание на продолжающейся дискуссии на тему авторского права: с одной стороны произведения, защищенные авторским правом, используются для обучения моделей ИИ, а с другой стороны сами модели способны генерировать реалистичный контент. Контент, генерируемый этими системами ИИ, может бросить вызов традиционным понятиям авторства и оригинальности.

Однако то, каким образом генеративный ИИ затрагивает интеллектуальную собственность (ИС), – гораздо более широкий вопрос, и стоит иметь в виду весь спектр прав ИС.

 Генеративный ИИ вызывает бурные дискуссии о возможном нарушении авторских прав в связи с включением охраняемых произведений в обучающие данные и тем, что создаваемый ИИ продукт охраняется авторским правом. Однако важно понимать, что вопросы ИС, связанные с генеративным ИИ, не ограничиваются авторским правом.



Генеративные модели ИИ

Очень важную роль играют патенты и коммерческая тайна, особенно в том, что касается охраны моделей ИИ, а также использования систем с открытым исходным кодом.

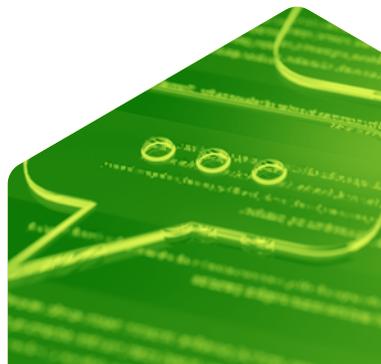
Патентной охране подлежат новые и обладающие признаками изобретения модели ИИ, алгоритмы и методы обучения. Однако патентоспособность систем ИИ различается в разных юрисдикциях, что создает проблему разграничения между патентоспособными системами ИИ и непатентоспособными абстрактными концепциями, реализуемыми с помощью стандартных компьютерных технологий. Кроме того, некоторые компании предпочитают открыто делиться необученными моделями, даже имея патенты на лежащие в основе этих моделей технологии.

В некоторых юрисдикциях авторско-правовая охрана исходного кода моделей ИИ и защита ключевых компонентов методов с помощью коммерческой тайны могут обеспечивать немедленную и более надежную охрану по сравнению с патентами.

 **Оптимальный способ обеспечить охрану генеративных моделей ИИ с помощью ИС зависит от нескольких факторов, включая характер модели ИИ, юрисдикцию и конкретные элементы, подлежащие охране. Патенты, коммерческая тайна и авторское право – все это играет определенную роль.**

Изобретательство с помощью генеративного ИИ

Генеративный ИИ потенциально способен комбинировать наборы знаний способами, которые людям могут показаться неочевидными, и, следовательно, способствовать инновационным прорывам. Он способен анализировать массивы данных большого объема, выявлять закономерности и предлагать решения, что способствует более эффективному решению задач и разработке инновационных методов.





Использование ИИ в процессе создания изобретений может дополнить инновационный потенциал человека, генерируя свежие идеи и решения. Люди по-прежнему играют решающую роль в процессе изобретательства, определяя проблемы, ставя цели и решая, как применять на практике идеи, генерируемые с помощью ИИ.

Вместе с тем использование генеративного ИИ в изобретательской деятельности может повысить планку изобретательского уровня, что потенциально может затруднить получение патентов на такие изобретения. В свою очередь изобретатели могут прибегнуть к защите своих инноваций с помощью коммерческой тайны, что потенциально может воспрепятствовать дальнейшим инновациям.

Вводимые данные для генеративного ИИ

Генеративные модели ИИ требуют большого объема обучающих данных. Например, исходная модель GPT-3 была обучена на 570 гигабайтах текста. В некоторых случаях части этих наборов обучающих данных были взяты из интернета и содержат текст и изображения, охраняемые авторским правом. В этой связи возникли вопросы о возможном нарушении авторских прав.

Разработчики генеративных моделей ИИ утверждают, что модели не хранят постоянную копию какой-либо работы, охраняемой авторским правом, а просто генерируют взвешенные значения в своих нейронных сетях. Таким образом, использование обучающих данных аналогично тому, как человек обучается, читая книги или осматривая произведения искусства. Они утверждают, что обучающие модели подпадают под действие существующих ограничений и исключений из авторского права, поскольку они не копируют обучающие данные. Большие наборы обучающих данных также гарантируют, что модели ИИ воспроизводят то, что создано человеком, и снижают риск систематической ошибки. Наконец, они говорят, что генеративный ИИ не заменяет собой творческую деятельность человека, а дополняет то, что создают люди, и служит дополнительным инструментом, способствующим появлению новых форм творчества.

Владельцы авторских прав, с другой стороны, утверждают, что несанкционированное использование их работ в качестве части наборов обучающих данных представляет собой нарушение этих прав.



Владельцами авторских прав подано несколько судебных исков против разработчиков ИИ, однако в ближайшем будущем вряд ли будут получены однозначные ответы на эти сложные юридические вопросы. Точно так же не существует простого способа оценить, насколько та или иная работа способствовала обучению модели ИИ и каким образом можно компенсировать ущерб, причиненный владельцу авторских прав. Кроме того, во многих юрисдикциях регистрация авторских прав не является обязательной, что иногда затрудняет идентификацию произведений, охраняемых авторским правом.

 Существуют противоречащие друг другу точки зрения относительно баланса между интересами владельцев авторских прав и тем, что разработчикам генеративного ИИ необходимо использовать произведения, охраняемые авторским правом, в качестве части своих наборов обучающих данных.

Результаты использования генеративного ИИ

Генеративный ИИ способен создавать широкий спектр творческих результатов, и сейчас активно обсуждается вопрос о том, должны ли эти результаты, создаваемые с помощью ИИ, охраняться авторским правом. Учитывая продуктивность генеративного ИИ, может появиться большое количество контента, разного по качеству и степени креативности, что ставит под вопрос устоявшиеся юридические понятия оригинальности, авторства и собственности.

 Авторское право было создано как право, ориентированное на человека, и во многих странах только люди могут быть авторами и, следовательно, владельцами авторских прав. Однако в настоящее время эта концепция оспаривается в некоторых юрисдикциях.

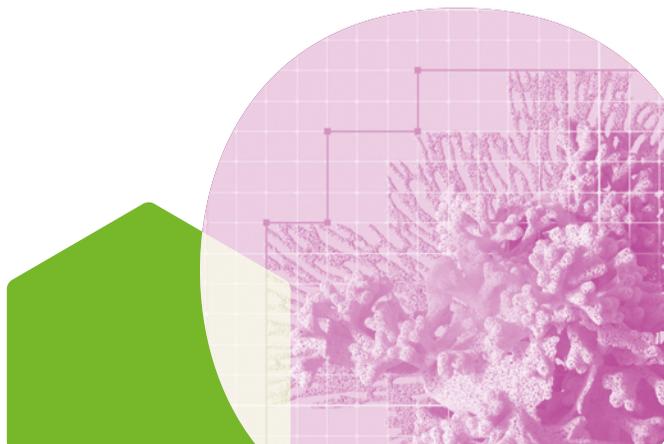
Например, некоторые заявители в Соединенных Штатах испрашивали охрану контента, созданного исключительно с помощью ИИ, но им было отказано. Совсем недавно интернет-суд Пекина постановил, что изображения, созданные с помощью Stable Diffusion, считаются оригинальными произведениями, поскольку многие подсказки, сделанные человеком, свидетельствуют о том, что эти изображения в достаточной степени являются результатом творческой деятельности человека.

В то же время генеративный ИИ может быть использован в качестве инструмента, с помощью которого создатели-люди могут открывать новые области творчества и создавать новые произведения. Если человек использует генеративный ИИ в качестве инструмента, но сам вносит значительный вклад в создание чего-либо, то результат работы потенциально может подлежать охране авторским правом. Остается открытым вопрос о том, как измерить такой вклад и установить какое-то пороговое значение в этом смысле.

 В мире, где ИИ играет существенную роль в творческом процессе, продолжают активные дискуссии о значении оригинальности.

Практические примеры

Генеративный ИИ активно развивается и используется для многих целей. Это разнообразие было продемонстрировано в ходе восьмого раунда Дискуссии ВОИС. Обсуждались самые разные примеры – от хореографических постановок до написания новостных статей. Еще один пример показал, как генеративный ИИ помогает представителям обособленных групп населения, таким как глухие и слабослышащие люди, распознавать язык жестов. Эти примеры также демонстрируют потенциально широкий спектр проблем, которые могут возникнуть в этой связи с точки зрения ИС.



Примеры

Открытые генеративные модели ИИ

Дополненные ИИ производство модных изделий и телепроизводство

Автоматическая обработка текста

Создание музыкальных произведений

Проблемы/вопросы с точки зрения ИС

Доступ к качественным обучающим данным имеет решающее значение для повышения безопасности использования ИИ и снижения вероятности систематических ошибок. Сложно ориентироваться в различиях между юрисдикциями в плане ограничений и исключений из авторского права и добросовестного использования.

Трудно определить, кто является владельцем контента, созданного с помощью ИИ. Из-за недостаточного разнообразия обучающих данных возможны незаконное присвоение и искажение традиционных выражений культуры.

Получение обучающих данных и создание обученных моделей требуют значительных затрат и времени.

Если модели не могут охраняться правами ИС, существует риск обратной разработки конкурирующих систем клиентами или конкурентами.

Соблюдение правил ИС в отношении обучающих данных является вопросом, беспокоящим небольшие фирмы, поскольку крупные игроки могут позволить себе действовать гораздо более рискованно.

Произведения, созданные с помощью ИИ, не подлежат охране авторским правом в большинстве юрисдикций, в которых рассматривался этот вопрос, что создает проблемы для радио и потоковых сервисов. Несанкционированное использование песен, музыки и текстов для машинного обучения вызывает обеспокоенность.

Авторы и исполнители должны иметь возможность запрещать использовать свой контент для обучения ИИ.



Программное обеспечение для распознавания языка жестов и перевода

Охрана моделей ИИ может помешать дальнейшим исследованиям и ограничить доступ к этим моделям.

Персонализированная телевизионная реклама

Авторские права на материалы и право собственности на них, особенно в случаях многочисленных участников, остаются неопределенными.

Поставщики стоковых изображений

Визуальные данные во многих случаях охраняются авторским правом, что требует получения разрешения на их использование для машинного обучения.

Эти примеры иллюстрируют потенциальные противоречия между интересами разработчиков генеративных моделей ИИ, создателей-людей и владельцев авторских прав. Технологические компании требуют, чтобы данные стали более доступными, в то время как правообладатели считают, что их работы не должны использоваться без согласия и справедливого вознаграждения.

Стратегии снижения риска

Снижение рисков, связанных с использованием генеративного ИИ, имеет решающее значение для компаний, выходящих на рынок в этой динамично развивающейся области.

 Компании могут принять ряд стратегических мер, позволяющих использовать потенциал ИИ и при этом сводящих к минимуму негативные последствия.

Для разработчиков ИИ

Очень важно проявлять должную осмотрительность. Компаниям следует рассмотреть возможность заключения всеобъемлющих контрактов как с поставщиками услуг, так и с поставщиками данных, поскольку это имеет решающее значение для защиты от потенциальных юридических обязательств, которые могут возникнуть в процессе разработки ИИ. Например, использование лицензионных соглашений в отношении данных, получаемых от третьих сторон, при обучении моделей ИИ может помочь обеспечить соблюдение прав ИС. Компании также могут рассмотреть возможность заключения контрактов и внедрения технологий, которые помогут защитить их системы ИИ и конфиденциальные данные от потенциальных угроз и обратной разработки.

Для пользователей моделей ИИ

Компании должны иметь четкое представление об условиях использования инструментов генеративного ИИ. Это может помочь компаниям разобраться в потенциальных лицензионных ограничениях и понимать, в чем заключается их юридическая ответственность.

Для владельцев авторских прав

Возможно, имеет смысл прямо указать, что охраняемое авторским правом произведение не может быть использовано для обучения моделей ИИ, или установить условия, при которых такие произведения могут использоваться для этих целей. Кроме того, применение технических мер защиты может обеспечить дополнительную защиту работ, охраняемых авторским правом, от несанкционированного использования.



Дополнительные материалы

Дискуссия ВОИС по вопросам ИС и передовых технологий является главным глобальным форумом для обсуждения и обмена знаниями о влиянии передовых технологий, включая ИИ, на ИС между всеми заинтересованными сторонами.

Главными темами обсуждения в ходе восьмого раунда Дискуссии ВОИС были генеративный ИИ и ИС, что было призвано помочь тем, кто определяет политику, понять, какая политика может проводиться в этом отношении. Более подробную информацию о восьмом раунде Дискуссии ВОИС, включая программу, презентации и веб-трансляцию, можно найти на соответствующей странице веб-сайта ВОИС.

Более подробную информацию об ИС и передовых технологиях см. на веб-сайте ВОИС: www.wipo.int/ai

Дальнейшие действия

Для получения информации о следующем раунде Дискуссии ВОИС подпишитесь на информационный бюллетень Отдела ИС передовых технологий.



