

П. С. Гуляева,
младший научный сотрудник,
Казанский инновационный университет
имени В. Г. Тимирязева

ЦИФРОВИЗАЦИЯ НОРМОТВОРЧЕСТВА В УСЛОВИЯХ СМЕНЫ ТЕХНОЛОГИЧЕСКОГО УКЛАДА

Аннотация. В рамках исследования рассмотрены теоретико-правовые аспекты применения генеративных языковых моделей в нормотворчестве. В частности, проведен эксперимент по разработке нормативных актов посредством популярных нейронных сетей ChatGPT и YandexLM, а именно: правила поведения на объектах метрополитена, проект акта об административном правонарушении и должностная инструкция младшего научного сотрудника. По итогам эксперимента был сформирован перечень проблем, возникающих при интеграции самообучающихся алгоритмов в нормотворческую деятельность.

Ключевые слова: нормотворчество, цифровизация, нейросеть, генеративная модель, YandexLM, ChatGPT, LLM, большие языковые модели

DIGITIZATION OF RULE-MAKING IN CONDITIONS OF CHANGE OF TECHNOLOGICAL STRUCTURE

Abstract. The study examines the theoretical and legal aspects of the use of generative language models in rule-making. In particular, an experiment was conducted to develop regulations using the popular neural networks ChatGPT and YandexLM, namely: rules of conduct at metro facilities, a draft act on an administrative offense and a job description for a junior researcher. Based on the results of the experiment, a list of problems that arise when integrating self-learning algorithms into rule-making activities was compiled.

Keywords: rule-making, digitalization, neural network, generative model, YandexLM, ChatGPT, LLM, Large Language Models

Изобретение колеса, парового двигателя, использование нефти как источника энергии повлекло не только смену экономической модели, но и серьезные социальные изменения. Технологическая трансформация XXI в. фундаментально отличается от предыдущих: инструментом и продуктом в новых условиях стала информация.

До 2020 г. технологии искусственного интеллекта и машинного обучения продемонстрировали разнообразный функционал с высокой степенью автоматизации, например, робот-поэт, робот-изобретатель, робот-судья. В правовой сфере можно отметить как зарубежные, так и отечественные проекты в рамках legaltech-тренда: Dashboard legal, Immediation, Jus Mundi, Botman.one и другие. Экспертные системы, чат-боты и конструкторы документов полностью или частично продублировали компетенции низкоквалифицированных юристов. При этом в правовой теории активизировалась дискуссия о базовых концептах, например, о гипотетической правосубъектности искусственного интеллекта и робототехники.

С 2021 г. генеративные языковые модели продемонстрировали принципиально новые подходы к цифровизации деятельности и работе с информацией. Генеративные нейронные сети можно классифицировать по типам следующим образом:

1) нейросети, формирующие по запросу оператора аудиовизуальный контент и способные генерировать и редактировать фото-, видеоматериалы (например, Midjourney, Kandinsky 2.0, Stable Diffusion);

2) нейросети, которые анализируют метаданные и данные, содержащие информационные следы пользователей (например, подбор целевой аудитории в рамках процесса подготовки рекламной компании в социальных сетях);

3) нейронные сети, применяемые в медицине, способны распознавать данные и обучаться на датасетах, собранных в разных городах страны, чтобы спрогнозировать состояние пациента более точно.

Популярность генеративных нейронных сетей обусловлена появлением, так называемых, Больших Языковых Моделей (Large Language Models – LLM), например, ChatGPT. Аналогичные сервисы разработаны компаниями Google, Apple, Yandex. Нейросети обучены подбирать информацию, тексты, формировать ответы на запросы оператора и участвовать в беседе в формате чата.

Генеративные нейронные сети обучены на огромных объемах данных, а LLM содержат, как правило, от 100 до 200 миллиардов параметров.

В процессе подготовки диссертационного исследования автора статьи был проведен эксперимент, в рамках которого была протестирована возможность применения сервисов ChatGPT и Yandex LM для автоматизации процесса нормотворческой деятельности. Были подготовлены несколько документов, а именно:

- правила поведения на объектах метрополитена;
- должностная инструкция младшего научного сотрудника;
- проект постановления об административном правонарушении.

Отечественный сервис Yandex LM уступает зарубежному ChatGPT по количеству параметров (100 миллиардов и 175 миллиардов соответственно). При этом нейросеть подходит для подготовки фрагментов текстов и характеризуется следующими параметрами:

- ограниченный объем итогового текста, который подходит для работы с короткими заметками либо при подготовке текста по частям;
- отсутствие понимания системой связи между запросами, а следовательно невозможность уточнить задачу после обращения;
- осмысленность результата приемлемая, но не превосходная;
- не подходит для работы в рамках высокой степени автоматизации;
- требуется не просто контроль результата, но и постоянное участие оператора.

В настоящее время специалисты Яндекс разрабатывают новую версию системы, которая обладает более совершенными техническими характеристиками.

Популярная зарубежная нейронная сеть продемонстрировала хорошие результаты, в частности:

- взаимосвязь запросов между собой;

- возможность уточнить параметры задачи как до запуска генерации, так и после него, возможность «подсказывать» системе и обращать внимание на логику посредством таких обращений, как «думаю последовательно» и подобных;
- сравнительно больший объем выдачи;
- понимание контекста достаточное для подготовки типовых документов на высоком уровне;
- качество работы приемлемое для тестирования высокоавтоматизированных сервисов, включая машиноисполняемые опции;
- к недостаткам нейросети можно отнести спонтанный переход с русского языка на иностранный.

Важно отметить, что указанные сервисы не обучены специально для выполнения правовых задач, но возможно подготовить нейронную сеть, адаптированную к работе с юридическими текстами.

В итоге к основным характеристикам генеративных языковых моделей можно отнести следующие признаки:

- 1) количество символов в итоговой выдаче ограничено возможностями конкретной генеративной языковой модели;
- 2) нейронная сеть функционирует на базе принципов и целей, заложенных разработчиками;
- 3) процесс принятия решения сервисом не прозрачен и неподконтролен – данный феномен называется проблемой «черного ящика»;
- 4) различаются системы с открытым и закрытым датасетом – в первом случае сервис учится на запросах пользователей по всему миру, во втором случае может учиться на обращениях только в рамках одной сессии, после завершения которой все удаляется;
- 5) генеративная модель не занимается поиском информации о конкретных людях, кроме сведений об общеизвестных личностях, которые были специально внесены в базу данных, например, ученый Стивен Хокинг;
- 6) уровень оригинальности текстов неодинаков;
- 7) в практике применения нейросетей в юриспруденции существуют кейсы, когда система «придумала» факты, например, по запросу судебного юриста сделала подборку выдуманных прецедентов, представив их как реальные – в итоге юрист был лишен права практиковать;
- 8) в зависимости от качества обучения нейронная сеть способна генерировать полноценные тексты, однако контроль человека, по мнению автора статьи, не только необходим, но и обязателен на современном этапе.

В процессе подготовки эксперимента были выявлены не только прикладные, но и фундаментальные противоречия:

- 1) понятийно-категориальный аспект связан с отсутствием терминологического аппарата, необходимого для разработки правового механизма и методологии применения нейросетей в нормотворческой деятельности;
- 2) проблема субъекта подразумевает сложности с формированием правового статуса нейросети и результатов ее деятельности, особенно в условиях высокой степени автоматизации нормотворчества на базе LLM;

3) деликтная проблема связана с предыдущей и характеризуется противоречиями при распределении ответственности за неблагоприятные последствия.

Помимо перечисленных противоречий можно отметить риски в сфере безопасности, а также этические, коммуникативные и методологические.

На основании вышеизложенного можно заключить, что в условиях развития цифровизации и, в частности, технологий машиночитаемого и машиноисполняемого права, эксперимент по интеграции инструментария больших языковых моделей в нормотворческую практику можно считать не только возможным, но и желательным. Применение генеративных нейронных сетей представляется верным в условиях цифровой трансформации и технологизации публичного и корпоративного управления. Самообучающиеся алгоритмы безусловно открывают новые горизонты в сфере цифровизации права, при этом провоцируют серьезные риски для правовой теории и практики.

Смена технологического уклада является объективным процессом и оказывает непосредственное влияние на правовую реальность, а также на индивида, общество в целом и государство. Разработка механизма применения самообучающихся алгоритмов в нормотворческой практике является в таком контексте естественным продолжением развития права в новых условиях цифровой трансформации всех сфер жизни. По мнению автора, игнорирование данных технологий и отказ от их применения в правовой сфере противоречит глобальным трендам и ведет к существенному отставанию в развитии цифровой экономики и социальных институтов.

Список литературы

1. Гуляева П. С. Медицинские нанороботы в фокусе права // *Journal of Digital Technologies and Law*. 2023. Т. 1, № 1. С. 89–122. EDN: WNRAOE.
2. Пашенцев Д. А., Залоило М. В., Дорская А. А. Смена технологических укладов и правовое развитие России. Москва: ИЗиСП: Норма: ИНФРА-М, 2021. 224 с.
3. Пономарева Е. В. Феномен квазисубъекта права: вопросы теории: монография / под редакцией С. И. Архипова. М.: Юрлитинформ, 2020. 154 с.
4. Цифровизация правотворчества: поиск новых решений: монография / отв. ред. Д. А. Пашенцев. М. Инфотропик Медиа, 2019.
5. Цифровизация правоприменения: поиск новых решений: монография / отв. ред. Д. А. Пашенцев. М. Инфотропик Медиа, 2022.
6. Alice Witt, Anna Huggins, Guido Governatori, and Joshua Buckley. Converting Copyright Legislation into Machine-Executable Code: Interpretation, Coding Validation and Legal Alignment. *Proceedings of the 18 International Conference on Artificial Intelligence and Law*. June 21–25, 2021. São Paulo, Brazil. Pp. 139–148.
7. Naman Jain, Arun Iyer, Suresh Parthasarathy, Sriram Rajamani, Rahul Sharma. Jigsaw: Large Language Models meet Program Synthesis. *Proceedings of the 44th International Conference on Software Engineering*, 2022. Pp. 1219–1231.