*Михаил Игоревич ГИЛИН*

# РАЗВИТИЕ ЦИФРОВИЗАЦИИ МЕДИАСФЕРЫ НА ПРИМЕРЕ CHATGPT: ВОЗМОЖНОСТИ И РИСКИ В ПРОЦЕССЕ УГЛУБЛЕНИЯ ТЕХНОЛОГИЧНОСТИ

*Статья посвящена применению искусственного интеллекта в медиасфере на примере чат-бота ChatGPT; рассматриваются тенденции видоизменения специальностей и развития технологии в медиапространстве.*

Ключевые слова: *ChatGPT, чат-бот, нейросети, копирайтинг, SEO*

*Mikhail I. GILIN*

# DEVELOPMENT OF MEDIA DIGITALIZATION ON THE EXAMPLE OF CHATGPT: OPPORTUNITIES AND RISKS APPEARING IN THE PROCESS OF DEEPENING OF TECHNOLOGICAL EFFECTIVENESS

*The article is devoted to application of artificial intelligence in media sphere, using chatbot ChatGPT as an example. It also examines the trends of specializations and development of technology in the media sphere.*

Key words*: ChatGPT, chatbot, neural networks, copywriting, SEO*

Появившийся в 2022 году в широком доступе чат-бот ChatGPT привнёс новые тенденции в развитие сразу нескольких взаимопроникающих сфер науки и бизнеса. Среди потенциально наиболее затронутых оказывается медиасфера – традиционно быстро меняющаяся и чутко реагирующая на движения социума и технологий обработки естественного языка.

Согласно информации на официальном сайте OpenAI, ChatGPT представляет собой созданную на основе языковой модели GPT-3,5 нейросеть со специализацией во взаимодействии с человеком посредством текстовых сообщений-команд в изолированном окне-интерфейсе на сайте разработчика. Чат-бот не имеет прямого выхода во всемирную паутину и оперирует заранее подготовленными, загруженными и проанализированными нейросетью данными из разных областей (см., например, цитату из ответа ChatGPT после соответствующего вопроса: *«As a language model AI, I don't have the ability to access live news updates or real-time information. However, I was trained on a diverse range of text from the internet, which includes news articles and other sources of information up until 2021»*. Перевод: *«Будучи*

*языковой моделью ИИ, я не имею возможности доступа к последним новостям и информации в реальном времени. Однако я обучался на разнообразных текстах из интернета, включающих новостные статьи и другие источники информации вплоть до 2021 года»*).

В рамках данного исследования было выявлено, что ChatGPT ознакомлен с широчайшим спектром тем: от популярных романов и текстов современных песен до стихотворений классиков и статей из научных журналов. Задачи для нейросети включали в себя составление списков вопросов с правильными ответами к рассказам Ray Bradbury: There Will Come Soft Rains и A Sound Of Thunder, Mark Twain: The Million Pound Banknote, Harlan Ellison: I Have No Mouth, and I Must Scream и Robert A. Heinlein: All You Zombies, а также генерацию собственных текстов песен и стихотворений в стилях фэнтезийной баллады, поп и рэп.

На основе терабайтов информации на всех основных мировых языках нейросеть становится способна формировать умения и проводить экспертизу в широком спектре вопросов, оказываясь полезным инструментом для лингвистов ряда специальностей. Чат-бот может быть использован главным образом для задач генерации текста как первого – наиболее трудозатратного и отнимающего время – этапа создания контента. По результатам данного исследования выявлено и будет продемонстрировано, что чат-бот умеет генерировать профессиональные статьи, заметки и новостные сводки. Однако в силу природы обучаемых нейросетей и своеобразного информационного пузыря, в котором они функционируют, использование в описанных целях лимитировано и требует интенсивной работы редактора и корректора. Отсутствие прямого доступа в интернет влечёт за собой неспособность чат-бота оперативно получать информацию. Учитывая ограничения в доступе к последним новостям, использование чат-бота видится более целесообразным при написании дайджестов и материалов по уже существующим в течение некоторого времени продуктам и услугам. Вторым критичным ограничением ChatGPT оказывается отсутствие модуля-надстройки для визуального распознавания деталей структур и объектов, о которых нейросеть получает специфические задания. При этом следующее поколение, GPT-4, находящееся на момент написания материала в разработке, по заявлениям разработчиков, лишена этого недостатка. С сайта OpenAI: *«GPT-4 can accept images as inputs and generate captions, classifications, and analyses»* (перевод: *GPT-4 может принимать изображения в*

*качестве вводных данных и генерировать субтитры, классификации и аналитические сводки*) (https://openai.com/product/gpt-4).

В рамках исследования возможностей ChatGPT проведён следующий эксперимент: чат-боту было поручено составить статью на неочевидную тему, а именно – об отдельно взятой модели наручных часов.

Оригинал: *«Rolex, the Swiss luxury watchmaker, has once again raised the bar for high-end timepieces with the release of their latest watch»;* перевод: *«Rolex, швейцарский производитель роскошных часов, в очередной раз поднял планку для элитных образцов, выпустив свою последнюю модель»*. Бросается в глаза пассаж *«элитные образцы»*, очевидным образом использованный из-за высокой частотности повторения в базе данных, на которой обучалась модель чат-бота.

Оригинал: *«The new Rolex watch combines timeless design with cutting-edge technology to offer a truly exceptional timepiece»*. Перевод: *«Новые часы Rolex сочетают в себе вневременной дизайн и передовые технологии, предлагая поистине исключительные часы»*. Словосочетания *«вневременной дизайн»* и *«передовые технологии»* также были использованы без оглядки на реально описываемые часы, с опорой на статистически вероятно более уместные.

В созданной ChatGPT статье текст был составлен следующим образом: использовались наиболее часто употребляемые в часовой индустрии слова и выражения – даже там, где они были неуместны и не соответствовали качествам действительно существующего предмета. Таким образом, при изучении текста получившейся статьи подтвердился принцип компиляции и предпочтения наиболее используемых терминов и структур, присущий другим лингвистическим нейросетям (Шаврина 2021: 119). ChatGPT, как и, например, системы машинного перевода, предпочитает использовать наиболее статистически вероятные слова и выражения. При этом провести самостоятельную проверку фактов система не способна в силу технологических ограничений.

В области медиа ChatGPT также подходит для составления слоганов и коротких маркетинговых текстов. Принимая во внимание определённую шаблонность, обыкновенно используемую авторами, эффективность работы ChatGPT для такого рода задач может быть оценена достаточно высоко. Описанная выше особенность принципов функционирование нейросетей, заключающаяся в использовании наиболее часто употребляемых единиц, в рамках написания

маркетинговых материалов играет на руку и позволяет сэкономить время и усилия лингвистов, задачей которых становится редактирование и фактчекинг.

В рамках проведённого исследования ChatGPT также был изучен в качестве ассистента для задач локализации, а именно – транскреации рифмованных массивов: стихов и песен. В качестве первого этапа эксперимента чат-боту было предложено создать текст к песне с целью исследования способности создавать рифмы согласно выбранному стилю и соблюдать ровный размер, см. цитату из составленной в рамках исследования песни:

*I'm the king of the ring, the master of rhymes,*
*With a flow so cold, it chills to the bones,*
*My punches hit harder than the toughest of chimes,*
*Leaving my opponents in the dust, all alone.*

Второй этап эксперимента заключался в переводе получившейся песни на русский язык с сохранением размера и порядка рифм. Цитата из получившейся в процессе перевода транскреации:

*Я – царь ринга, мастер рифм,*
*С теченьем так холодным, что морозит до костей,*
*Мои удары сильнее самых твёрдых дождей,*
*Оставляя моих противников в пыли, всех в одиночестве.*

Несмотря на сохранение смысла, чат-бот не справился с сохранением размера и частично – рифм. Однако очевидна попытка транскреации, при которой ChatGPT полностью изменил смысл второй части третьей строки: вместо громкого «перезвона/звона/удара» был использован образ «плотных дождей». По итогам эксперимента были сделаны выводы об ограниченной пригодности чат-бота для выполнения сложных переводческих задач. Получаемые рифмы были релевантны контексту, но не несли художественной ценности из-за их чрезмерной простоты.

Однако появление и развитие языковых продуктов среди высокоэффективных нейросетей манифестируют кризис отдельных специальностей как непосредственную реакцию на конкретную разработку и кризис лингвистики в общем – как отражение явной тенденции занятия всё большего количества ниш на стадии создания текстовых продуктов, будь то перевод, транскрипция или генерация. Уже сейчас специалисту-человеку проще найти себя в качестве постредактора или корректора, нежели переводчика или титровальщика (Нечаева, Светова 2018: 67). А потому высока

вероятность видоизменения или даже исчезновения специальностей лингвистов в медиасфере.

В качестве примера специальности в зоне риска оказывается рерайтер – специалист, привлекаемый для создания материалов на основе уже существующих текстов по заданным параметрам уникальности и с включением в тело текста ключевых слов и словосочетаний для повышения индексируемости в поисковых системах Интернета. Благодаря принципам своей работы ChatGPT способен не только мгновенно выполнить эту задачу, но даже предоставить несколько вариантов на выбор по запросу. Лингвисту остаётся проверить готовый текст на факты и выполнить минимальную корректуру. Учитывая, что такого рода тексты незамысловаты структурно и по смыслу, услуги человека по их написанию с нуля становятся неактуальными.

В ближайшем будущем видоизменяющейся специальностью может оказаться специальность копирайтера – автора статей для рекламы и других видов маркетинга. Перспектива переложить наиболее трудозатратный и длительный этап на «машину» более вероятна, нежели сохранение современной производственной цепочки. Наиболее очевиден переход от непосредственного написания статьи к постредактированию и фактчекингу, что потенциально уменьшит время от инициации до получения готового к публикации материала и снизит финансовые издержки издательств и заказчиков.

Однако говорить о полной замене человека компьютерной программой для задач создания текстового контента по меньшей мере преждевременно. С делегированием задач чат-боту проблема контроля качества готового продукта встаёт более остро. Если в производственной цепочке, предусматривающей только специалистов-людей, каждый последующий этап занимается в том числе контролем качества предыдущего, даже если это не декларируется, то при включении в производственную цепочку искусственного интеллекта, непредсказуемо часто генерирующего ложные факты, вероятность допущения ошибки увеличивается. Ещё более опасной видится ситуация несанкционированного использования нейросети при заявляемом «ручном труде» в силу необходимости меньшего контроля за автором-человеком, который, возможно, предпочёл бы более скупую и краткую сводку по предмету, а не заведомо ложную, но частотно более вероятную, суммаризацию.

С точки зрения долгосрочной перспективы использование чат-

ботов вместо собственноручного написания текстов может привести к всеобщему падению способности к сложному описанию логических цепочек и созданию текстовых продуктов без помощи генераторов текстов. С популяризацией технологии вероятно её проникновение в образование и деятельность школьников и студентов, поэтому образовательные институты уже сейчас рассматривают мероприятия для ограничения влияния и вероятности использования подобных инструментов в рамках специфических для формирования соответствующих навыков образовательных процессов. Из статьи ТАСС от 1 февраля 2023 года: *«РГГУ прокомментировал информацию о том, что студент вуза защитил написанный ChatGPT диплом. Университет предложил ограничить доступ к чат-боту в образовательных организациях»* (https://tass.ru/obschestvo/16937425).

Рассматривая другие вызовы, связанные с появлением технологии и продукта, важно упомянуть настоящую неопределённость в правовом поле. Из-за того, что перед обучением модели чат-бота из сети Интернет выбирается огромный массив данных, часть приобщённой информации неизбежно оказывается защищённой авторским правом. Насколько использование защищённых таким образом данных и компиляция на их основе станут нарушением авторских прав, на данный момент остаётся открытым вопросом.

Ещё одной проблемой, приобретающей актуальность именно на фоне развития информационных технологий, оказывается низкий уровень цифровой гигиены и общая неосведомлённость населения о принципах информационной безопасности (Гусева и др. 2017: 456). Так, безответственное использование чат-бота повлечёт утечку конфиденциальных данных в случае их наличия в размещаемых запросах к нейросети. По причине того, что центры обработки данных ChatGPT размещены за границей, вероятность нежелательных последствий находится на высоком уровне.

Появление новых языковых продуктов на основе нейросетей и неизбежная адаптация к ним – вопрос если не настоящего, то ближайшего будущего. Платой за ускорение производственных цепочек станет исчезновение или видоизменение всё большего количества специальностей, и специалисты должны быть готовы к развитию науки, технологий и – в последующем – рынков. Из вышеизложенного следует, что человек не будет исключён из процессов создания языковых продуктов, однако его итоговая роль предположительно сведётся к редактирующей и контролирующей функциям.

## Список литературы:

*Гусева А.Ю., Данфельд А., Рукавцова О.М.* «Глобальное информационное общество» и проблемы «информационной гигиены» // Контуры будущего: технологии и инновации в культурном контексте – 2017 – Режим доступа: https://elibrary.ru/download/elibrary_32791138_74279305.pdf

*Гусева В.Б.* Современные средства машинного перевода // Современные информационные технологии: проблемы и перспективы развития – 2017 – Режим доступа: https://elibrary.ru/download/elibrary_37298800_74297158.pdf

*Нечаева Н.В., Светова С.Ю.* Постредактирование машинного перевода как актуальное направление подготовки переводчиков в вузах // Вопросы методики преподавания в вузе – 2018 – Режим доступа: https://elibrary.ru/download/elibrary_36388371_63334176.pdf

*Шаврина Т.О.* О методах компьютерной лингвистики в оценке систем искусственного интеллекта // Вопросы языкознания – 2021 – Режим доступа: https://elibrary.ru/download/elibrary_47270094_34714791.pdf

---

*ГИЛИН Михаил Игоревич* – руководитель магистерской программы «Цифровая лингвистика и локализация» Национального исследовательского технологического университета «МИСИС» (г. Москва), *m.gilin@misis.ru*