RUDN Journal of Studies in Literature and Journalism
Вестник РУДН. Серия: Литературоведение. Журналистика
ISSN 2312-9220 (Print); ISSN 2312-9247 (Online)
2024 Vol. 29 No. 3 593–611
http://journals.rudn.ru/literary-criticism

МЕДИАПОЛИТИКА

MEDIA POLICY

DOI: 10.22363/2312-9220-2024-29-3-593-611
EDN: GICACA
УДК 327

Научная статья / Research article

Анализ текстов о спорном статусе Тайваня, созданных языковой моделью ChatGPT

А.Ю. Петухов[1,2], Д.И. Каминченко[2,3] ✉

[1] *Московский государственный университет имени им. М.В. Ломоносова, Москва, Российская Федерация*
[2] *Университет «Неймарк», Нижний Новгород, Российская Федерация*
[3] *Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Российская Федерация*
✉dmitkam@inbox.ru

Аннотация. Технологии искусственного интеллекта становятся все более распространенными в современном обществе. Не остается в стороне и сфера политики, где также рассматриваются и начинают применяться указанные технологии. Цель авторов – выявить особенности презентации текстов, генерируемых одним из наиболее популярных онлайн-сервисов, нейросетью ChatGPT, по конкретным ключевым текстовым запросам. Запросы связаны с двумя подходами к проблематике политического статуса Тайваня: как части Китайской Народной Республики или как независимого и суверенного территориального образования. В качестве базовых методов использованы контент-анализ количественного типа и фрейм-анализ. Информационный массив данных составили результаты, полученные по итогам двух текстовых запросов: «Taiwan as an independent entity in...» и «Taiwan as part of China in...» по четырем временны́м точкам (2010, 2015, 2020 и 2023 гг.), выбранным на основании принципа равной удаленности и периода проведения исследования (2023 г.). Общее количество слов для анализа – 1402. В результате исследования установлено, что по текстовому запросу, напрямую связанному с независимым и суверенным политическим статусом Тайваня, искусственный интеллект генерирует тексты бо́льшего объема, чем по текстовому запросу, относящемуся к утверждению о его территориальной принадлежности КНР. С точки зрения символических слов/выражений-маркеров, свидетельствующих о том или ином подходе к указанной политической проблематике, в генерируемом искусственным интеллектом контенте выявлено два момента. Первый – абсолютное преимущество слов/выражений-маркеров,

© Петухов А.Ю., Каминченко Д.И., 2024
This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

символически отражающих идею о независимом и суверенном политическом статусе Тайваня, второй – минимизация слов/выражений-маркеров, отражающих территориальную принадлежности Тайваня КНР. Обозначены возможные причины подобного распределения частоты появления в информационном массиве указанных слов/выражений-маркеров. Установлено наличие во всем информационном массиве как минимум трех распространенных фреймов, которые обозначены следующим образом: «Independent entity», «The part of China», «Tensions and confrontation».

Ключевые слова: ChatGPT, языковые модели, нейронные сети, искусственный интеллект, Тайвань, контент-анализ, Китайская Народная Республика

Вклад авторов. *А.Ю. Петухов* – разработка идеи исследования, анализ данных, написание и редактирование рукописи. *Д.И. Каминченко* – разработка идеи исследования, сбор исследовательских данных, анализ данных, написание и редактирование рукописи.

Заявление о конфликте интересов. Авторы заявляют об отсутствии конфликта интересов.

История статьи: поступила в редакцию 18 июня 2024 г.; отрецензирована 1 июля 2024 г.; принята к публикации 12 июля 2024 г.

Для цитирования: *Петухов А.Ю., Каминченко Д.И.* Анализ текстов о спорном статусе Тайваня, созданных языковой моделью ChatGPT // Вестник Российского университета дружбы народов. Серия: Литературоведение. Журналистика. 2024. Т. 29. № 3. С. 593–611. http://doi.org/10.22363/2312-9220-2024-29-3-593-611

Analysis of ChatGPT-generated texts on Taiwan's disputed status

Aleksandr Yu. Petukhov[1,2], Dmitriy I. Kaminchenko[2,3] ✉

[1]*Lomonosov Moscow State University, Moscow, Russian Federation*
[2]*Neimark University, Nizhny Novgorod, Russian Federation*
[3]*Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russian Federation*
✉dmitkam@inbox.ru

Abstract. Artificial intelligence technologies are gradually becoming more common in modern society. The political sphere is not left aside, where these technologies are also being considered and are beginning to be applied. The purpose of this research is to identify the features of the presentation of texts generated by one of the most popular online services, neural network ChatGPT, for key text queries associated with two approaches to the issue of the political status of Taiwan: considering Taiwan as part of the Chinese People's Republic or as an independent and sovereign territorial entity. Quantitative content analysis and frame analysis are used as basic methods. The information array of data consists of the results obtained from two text queries: "Taiwan as an independent entity in..." and "Taiwan as part of China in..." for four time points (2010, 2015, 2020 and 2023), selected based on the principle of equal distance and the period of the study (2023). The total number of words for analysis was 1402. The study found that for a text query directly related to the independent and sovereign political status of Taiwan, artificial intelligence generates larger texts than for a text query related to the statement of its territorial affiliation with the PRC. From the point

of view of symbolic words/expressions-markers, indicating a particular approach to the specified political issues, the content generated by artificial intelligence revealed the absolute advantage of words/expressions-markers, symbolically reflecting the idea of the independent and sovereign political status of Taiwan, and not its territorial affiliation with the PRC. Possible reasons for such a distribution of the frequency of occurrence of the specified words/expressions of markers in the information array are indicated. The presence of at least three common frames in the entire information array has been established, which are designated as follows: "Independent entity", "The part of China" and "Tensions and confrontation".

Key words: ChatGPT, language models, neural networks, artificial intelligence, Taiwan, content analysis, People's Republic of China

Author's contribution. *Aleksandr Yu. Petukhov* – development of the research idea, data analysis, manuscript writing and editing. *Dmitry I. Kaminchenko* – development of the research idea, research data collection, data analysis, manuscript writing and editing.

Conflicts of interest. The authors declare that there is no conflict of interest.

Article history: submitted June 18, 2024; revised July 1, 2024; accepted July 12, 2024.

For citation: Petukhov, A.Yu., & Kaminchenko, D.I. (2024). Analysis of ChatGPT-generated texts on Taiwan's disputed status. *RUDN Journal of Studies in Literature and Journalism, 29*(3), 593–611. (In Russ.) http://doi.org/10.22363/2312-9220-2024-29-3-593-611

Введение

Технологии искусственного интеллекта активно внедряются в различные сферы общественного функционирования. Инструменты, генерирующие из огромного массива данных тексты и фотоизображения, становятся крайне популярными среди разных представителей общества. Один из наиболее известных из них – ChatGPT. Однако параллельно стали появляться и множественные аналоги, в том числе российские, например, YandexGPT, встроенный в голосовой помощник «Алиса», GigaChat – российский аналог ChatGPT от «Сбера» и другие.

Появление подобных инструментов, во многом обусловленное развитием технологий и алгоритмов, основанных на нейронных сетях, инициировало большое количество вопросов. Действительно, языковые модели могут существенно облегчить решение ряда задач, но при этом и создать характерные проблемы, в том числе – непредвзятости информации, выдаваемой с помощью подобных алгоритмов (Одноконный, Рогач, 2023; Baum, Villasenor, 2023). Ведь данные инструменты используются не только для относительно нейтральных запросов, например, поиска информации о конкретном животном или географическом объекте, но также и о куда более субъективных вещах[1].

Для большинства людей такой вопрос может показаться странным, так как сбор и обработку данных осуществляет алгоритм, не обладающий

[1] ChatGPT breaks its own rules on political messages // The Washington Post. August 28, 2023. https://www.washingtonpost.com/technology/2023/08/28/ai-2024-election-campaigns-disinformation-ads/

(чего напрасно опасаются некоторые группы граждан) собственным самосознанием и мнением (Беляков, Максименко, 2023). Однако, очевидно, что содержание текстов будет зависеть от двух основных факторов:

— специфика алгоритма и его настройка, на которой основан инструмент (т. е. придает ли он бóльшее значение тем или иным источникам, отдельным маркерам, словам, позициям и т. д.);

— массив текста, используемый для обработки и обучения (насколько он широк, включает ли в себя достаточное количество источников с различными позициями и т. д.).

Очевидно, что даже без специализированной подстройки алгоритмов, если использовать в качестве массива данных исключительно, допустим, западноевропейские источники информации (СМИ, базы данных, государственные сайты), то результат запроса на их основе будет кардинально отличаться от результата при преимущественном использовании ресурсов Китая, России или стран арабского мира. Это связано с разной интерпретацией фактов, трактовок, элементов пропаганды, культурных различий и многого другого (Subbotkin, 2017; Lidgren, 2019; Петухов и др., 2023).

Результаты генерации подобных инструментов уже используются при написании статей в СМИ, анализе исторических данных, в том числе исследователями, а не только рядовыми пользователями. Скорее всего, частота использования будет расти, а, следовательно, с помощью обозначенных инструментов возрастет и масштаб влияния со стороны различных субъектов политики на политические взгляды индивидов, их культурную и национальную идентичность и т. д.

Кроме того, ввиду общей специфики и архитектуры алгоритмов, ChatGPT, по сути, должен являться неким общим «сумматором» накопленных массивов данных в сети Интернет по запрошенной тематике, т. е. по результатам генерации текстов можно также делать осторожный вывод об общих тенденциях в мировых СМИ, базах данных, некоторых социальных сетях.

Указанные аспекты на данный момент не были проанализированы в достаточной мере: на момент написания статьи в международных базах цитирования практически отсутствуют исследования, посвященные выбранной тематике. Безусловно, это в бóльшей степени связано с крайне быстрым развитием технологий и специфическими задержками с публикацией исследований и их проведением.

Данная работа направлена на изучение специфики генерируемых текстов по конкретной проблематике (политический статус Тайваня в современном мире) одним из подобных ресурсов – ChatGPT. Вопрос политического статуса выбран не случайно: это весьма значимая геополитическая проблема, которая характеризуется наличием двух непримиримых сторон, возможностью перерастания в военный конфликт, а также высокой значимостью всех сторон для мировой экономики и политики (Chou,

2018). Актуальность данной темы отражается и в сохраняющемся в последние годы интересе к ее рассмотрению в научном сообществе (Конуров, 2020; Дикарев, Лукин, 2021; Гудев, Мишин, 2022; Пузыня, 2022; Фадеева, 2023).

Цель исследования заключается в выявлении особенностей презентации соответствующих текстов по ключевым текстовым запросам, связанным с двумя подходами к проблематике работы: рассмотрению Тайваня как части Китайской Народной Республики или как независимого и суверенного территориального образования.

Методы и материалы

Подбор исследовательских методов и порядок их применения обусловлены основной целью работы. В качестве базовых методов используются контент-анализ количественного типа и фрейм-анализ. Количественный контент-анализ «может быть представлен как систематизированное изучение содержания письменного или устного текста с фиксацией наиболее часто повторяющихся в нем словосочетаний или сюжетов» (Боришполец, 2012, с. 62). Такой анализ позволит выявить наличие/отсутствие ключевых слов-маркеров, содержащих в себе символическую коннотацию применительно к выбранной проблематике и частоту их использования. В данном случае целесообразно прибегнуть к ненаправленному контент-анализу, который позволит выявить подобные слова-маркеры непосредственно в ходе проведения само*го* текстового анализа.

Особое внимание должно быть уделено сбору и формированию информационного массива данных, так как необходимо подобрать поисковый текстовый запрос, релевантный двум подходам к проблематике политического статуса Тайваня. В состав текстового запроса должны быть включены как минимум две части: содержательно-смысловая и временна́я.

Первая будет по характерным маркерам определять позицию, доминирующую в конкретном текстовом результате информационного запроса (содержащую позиционирование Тайваня как независимого государства (Жук, 2023) или нет). Вторая часть, временн*а*я, будет включать конкретный год, чтобы можно было отследить динамику результата и минимизировать шанс ошибки за счет выборки (например, если случайным образом в какой-то период времени одна из точек зрения аномально превалировала).

В результате, после проверки различных вариантов, были сформулированы два шаблона текстовых запросов:

— «Taiwan as an independent entity in...»;

— «Taiwan as part of China in...».

За основу были взяты четыре временн*ы*х точки: 2010, 2015, 2020 и 2023 гг. Итоговые текстовые запросы выглядели следующим образом: «Taiwan as an independent entity in 2010», «Taiwan as part of China in 2010»

и т. п. Временны́е точки выбраны на основе принципа равноудаленности: они включают в себя одинаковые интервалы, за исключением 2023 г. (год проведения исследования). По ходу работы выяснилось, что алгоритмы инструмента позволяют генерировать тексты за период до 2021 г. В случае текстов, датированных более поздними годами, ситуация сложнее. Они представляют собой некие прогнозы, формируемые искусственным интеллектом. Возможно, это результат смешения годов исследований, т. е. алгоритмы могут выбрать текст 2022 г. для анализа по 2010 г. (так как в нем писалось о ситуации в 2010 г.) или же текст 2010 г. для анализа по 2022 г. (так как в нем содержался прогноз на 2022 г.). Конечно, это несколько размывает временны́е границы и точность динамики, но, поскольку основная цель исследования иная, данные искажения не снижают релевантности выводов.

Проведение контент-анализа позволит установить количество слов-маркеров, представленных в генерируемых ИИ текстах, а сравнительный анализ по разным годам даст возможность отследить динамику изменения представленности конкретных слов и специфики выражаемого с их помощью смысла.

Важную роль в проведении текущего исследования играет и применение метода фрейм-анализа. Термин «фрейм», как пишет В.С. Вахштайн[2], является собирательным обозначением контекста» (Вахштайн[3], 2011, с.41). В научных исследованиях встречаются разнообразные примеры задействования указанного метода. В частности, Н. Раушер и М. Шефер в одной из своих недавних работ на примере анализа избирательной кампании Э. Уоррен, участницы праймериз Демократической партии США 2019–2020 гг., рассматривают вопрос о том, какие нарративы о неравенстве используются в (прогрессивном) политическом дискурсе в Соединенных Штатах («(progressiven) politischen Diskurs der Vereinigten Staaten») и как они передаются и оцениваются с использованием фреймов (Rauscher, Schäfer, 2022).

Применение фрейм-анализа в текущей работе позволит выделить конкретные символико-смысловые/тематические единицы-фреймы и определить их популярность в текстах, полученных по различным запросам. Естественно, часть установленных фреймов будет в некоторой степени перекликаться со сформированными текстовыми запросами для ChatGPT, но их выделение является крайне важным с точки зрения популярности и наличия/отсутствия преобладания одного из двух подходов к обозначенной проблематике. В свою очередь, информация о выявленных фреймах будет дополнена конкретными словами, словосочетаниями и выражениями, подкрепляющих тот или иной фрейм.

[2] Внесен Министерством юстиции РФ в реестр СМИ, выполняющих функции иностранного агента.

[3] Внесен Министерством юстиции РФ в реестр СМИ, выполняющих функции иностранного агента.

Результаты и обсуждение

Информационный массив данных составил шесть сгенерированных искусственным интеллектом текстов по обозначенным текстовым запросам, общее количество слов для анализа – 1402. Распределение общего объема слов по годам и разным текстовым запросам отображено на рис. 1.

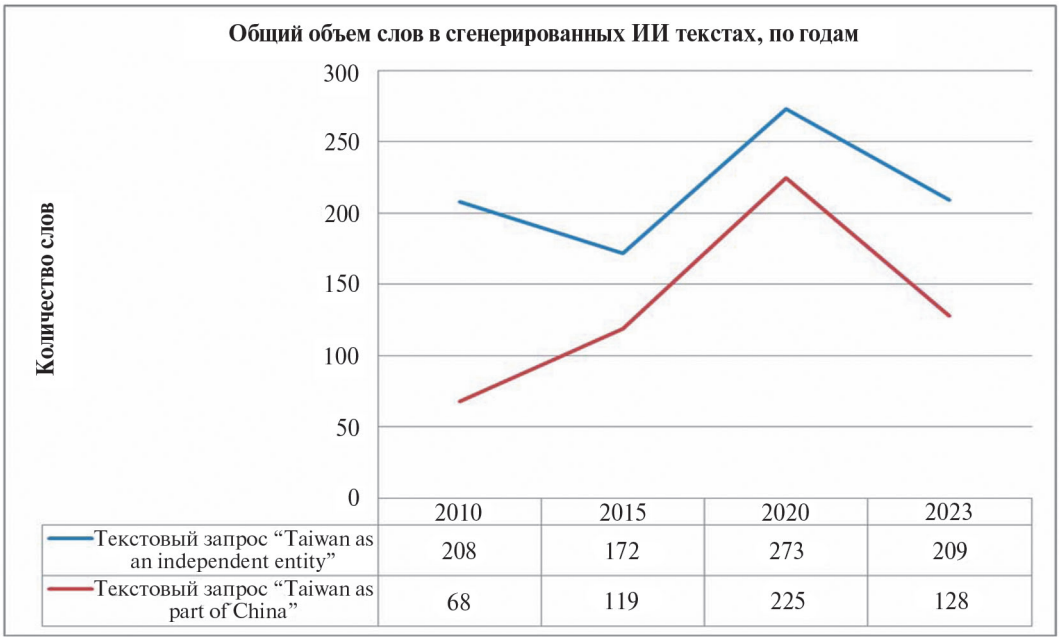


Рис. 1. Общий объем слов в сгенерированных ИИ текстах, по годам

Источник: составлено авторами на основе данных ChatGPT

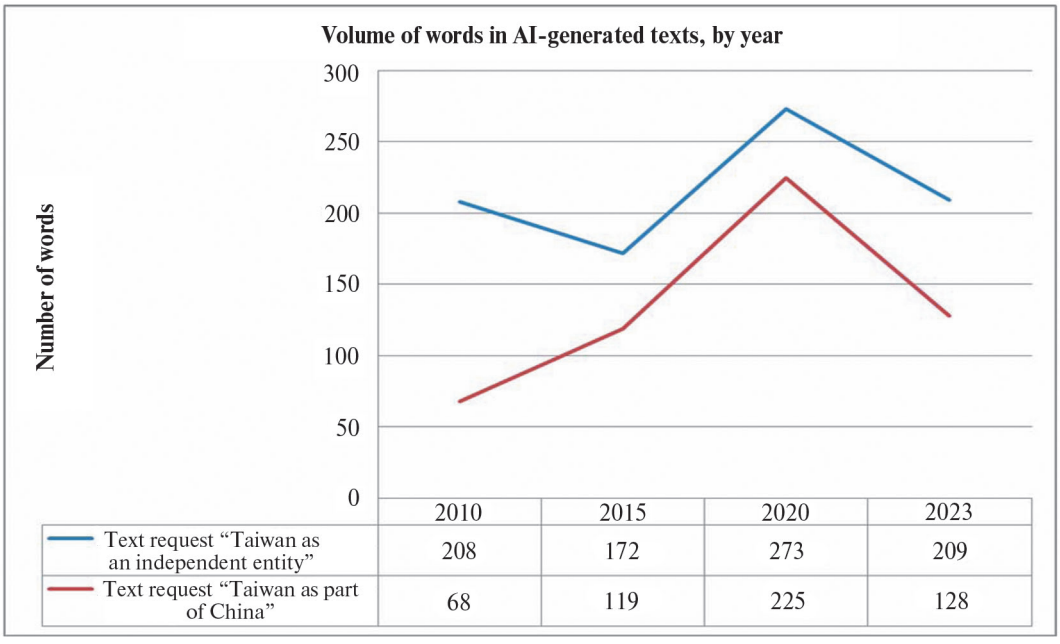


Figure 1. Volume of words in AI-generated texts, by year

Source: compiled by the authors based on ChatGPT data.

Уточним, что формулировки текстовых запросов различаются в зависимости от года своей временно́й составляющей, например, за 2010 г. использовался такой запрос, как «Taiwan as part of China in 2010», а за 2015 – «Taiwan as part of China in 2015» и т. п. Согласно данным таблицы, по запросу «Taiwan as an independent entity» ИИ генерирует тексты б*о*льшего объема, чем по запросу «Taiwan as part of China».

Максимальный по объему текст по запросу «Taiwan as an independent entity» генерируется ИИ в 2020 г., минимальный – в 2015 г. Максимальный по объему текст по запросу «Taiwan as part of China» генерируется ИИ также в 2020 г., а минимальный – в 2010 г.

Линии тренда практически параллельны в случае двух текстовых запросов в два периода: с 2015 по 2020 г. и с 2020 по 2023 г. В период с 2010 по 2015 г. наблюдается обратное изменение обеих линий (изменение объема слов сгенерированных текстов).

Проводя контент-анализ, мы выделили слова-маркеры, использование которых в представленном в информационном массиве контексте говорит либо о независимости и суверенитете Тайваня как территориальной единицы, либо как о части КНР (Сунь, 2018; Шаров, 2018). В первом случае речь идет о следующих словах/выражениях-маркерах: «independent», «own» «international», «state», «diplomatic», «separate», «country», «nation», «foreign», «Republic of China», «sovereign», «Chinese Teipei», «WTO», «UN». Всего в сгенерированных текстах по обоим типам запросов и во все годы (в соответствующем смысловом контексте) данные слова использовались ИИ 88 раз. Во втором случае речь идет о таких словах/выражениях-маркерах, как: «part/province of its/China territory», «sovereignty», «control», «One China Policy», «representative of China». Всего в сгенерированных текстах по обоим типам запросов и во все годы (в соответствующем смысловом контексте) данные слова использовались ИИ 18 раз. Распределение частоты слов/выражений-маркеров по обоим типам запросов по разным годам представлено на рис. 2.

Как видно из диаграммы и таблицы, ИИ намного чаще генерирует слова/выражения-маркеры, которые говорят о независимом и суверенном статусе Тайваня как территориальной единицы, чем как о части КНР, причем наблюдается нелинейная динамика изменения по годам с пиковым значением в 2020 г. и с минимальным значением в 2023 г. Частота встречаемости слов/выражений-маркеров, отражающих территориальную принадлежность Китаю, оказалась одинаковой на протяжении сразу трех временных промежутков, и только в 2023 г. обозначенные слова используются значительно меньшее количество раз. Распределение конкретных слов/выражений-маркеров по обозначенным текстовым запросам по годам представлено на рис. 3.

В случае со словами/выражениями-маркерами, отражающими независимый и суверенный статус Тайваня, наблюдается более активное их генерирование в текстах по запросам «Taiwan as an independent entity», чем



Рис. 2. Общее количество слов/выражений-маркеров в сгенерированных ИИ текстах по запросам обоих типов, по годам

Источник: составлено авторами на основе данных ChatGPT.

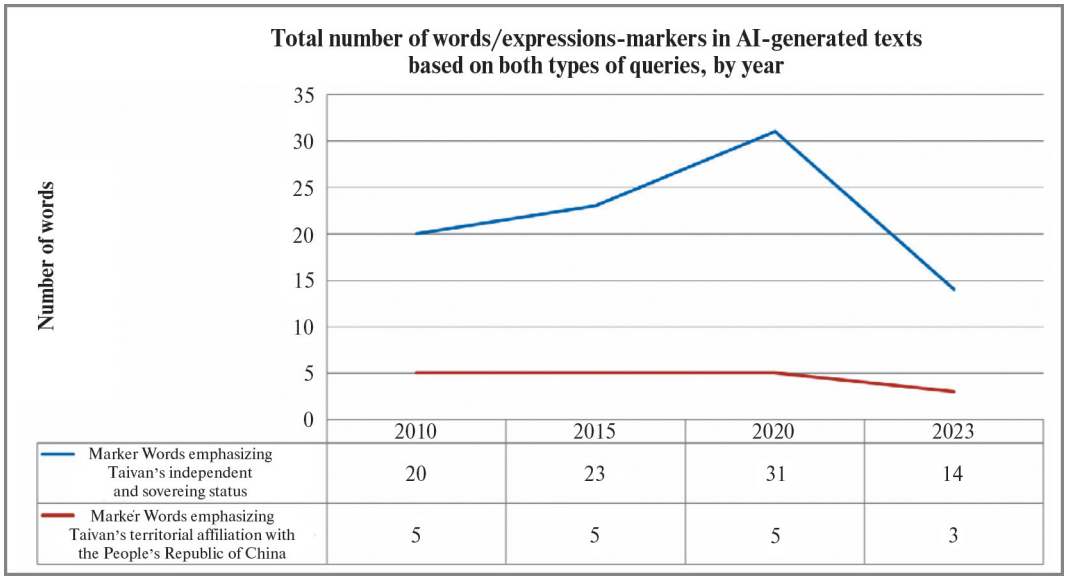


Figure 2. Total number of words/expressions-markers in AI-generated texts based on both types of queries, by year

Source: compiled by the authors based on ChatGPT data.

по запросам «Taiwan as part of China», что вполне объяснимо, причем наблюдается и схожая динамика изменений в два временны*х* периода, с 2015 по 2020 г. и с 2020 по 2023 г. Относительно слов/выражений-маркеров, отражающих территориальную принадлежность Тайваня КНР, ситуация несколько иная: общая частота генерирования подобных слов по обоим текстовых запросам одинаковая (по 9 раз) с некоторыми расхождениями по

разным годам наблюдений. Впрочем, необходимо обратить внимание на небольшую частоту генерирования данных слов в целом по разным годам.

По результатам проведения фрейм-анализа установлено наличие во всем информационном массиве как минимум трех распространенных фреймов, которые условно можно обозначить следующим образом: «Independent entity», «The part of China» и «Tensions and confrontation». О наличии выделенных фреймов свидетельствуют слова и выражения, представленные в сгенерированных текстах. В частности, фрейм «Independent entity» составляют следующие лексемы, словосочетания и выражения: «Republic of China» («ROC»), «considers itself a sovereign state», «own government», «own military», «its de facto independence», «own constitution», «has own political system» и др.

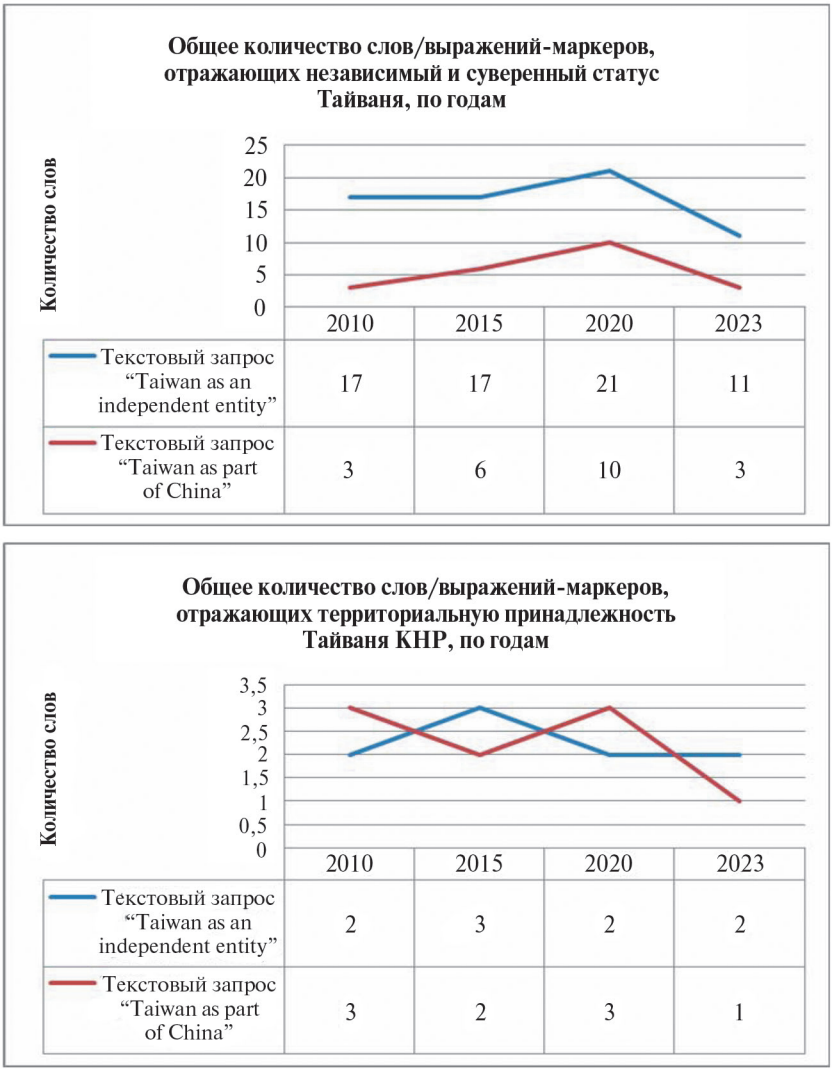


Рис. 3. Общее количество слов/выражений-маркеров двух тематических направлений в сгенерированных ИИ текстах по запросам обоих типов, по годам

Источник: составлено авторами на основе данных ChatGPT.

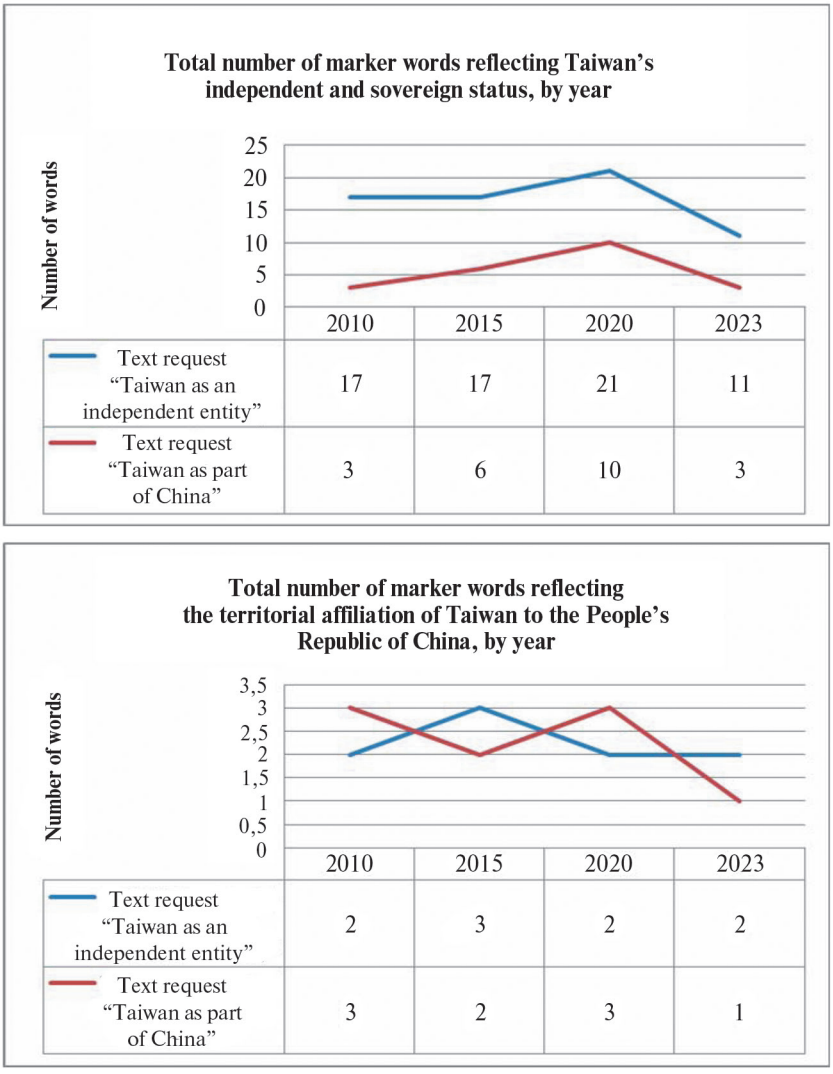


Figure 3. Total number of marker words of two thematic areas in AI-generated texts based on queries of both types, by year

Source: compiled by the authors based on ChatGPT data.

Содержание фрейма «Part of China» составляют следующие языковые единицы: «not recognized as an independent entity», «to be a part of its territory», «not as a fully recognized independent state», «not recognized by the United Nations as an independent country», «the People's Republic of China (PRC) considers Taiwan to be a part of its territory» и др.

Наличие фрейма «Tensions and confrontation» подтверждается следующими сгенерированными словами и выражениями: «has not renounced the use of force to bring it under its control», «ongoing tensions between Taiwan and China», «pressure from China», «China tries to isolate Taiwan diplomatically», «China tries to limit its international recognition», «the People's Republic of China (PRC) strongly opposes any moves toward Taiwanese independence», «ongoing pressure from the PRC to reintegrate with the mainland» и др.

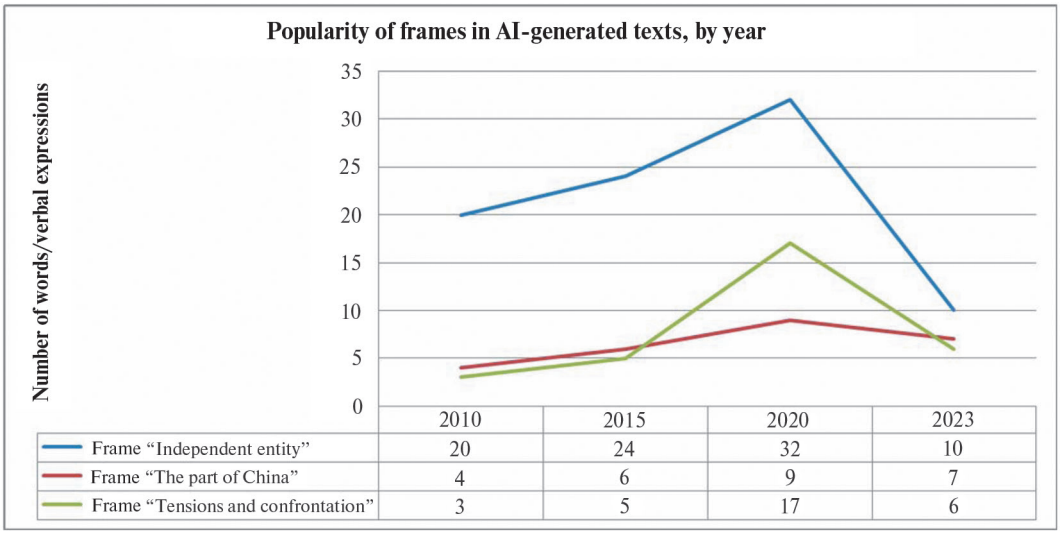
На основании подсчета количества слов и выражений, составляющих тот или иной фрейм, можно определить степень популярности фреймов по отношению друг к другу по годам. Данная статистика приведена на рис. 4.



	2010	2015	2020	2023
Фрейм “Independent entity”	20	24	32	10
Фрейм “The part of China”	4	6	9	7
Фрейм “Tensions and confrontation”	3	5	17	6

Рис. 4. Популярность фреймов в сгенерированных ИИ текстах, по годам

Источник: составлено авторами на основе данных ChatGPT.



	2010	2015	2020	2023
Frame “Independent entity”	20	24	32	10
Frame “The part of China”	4	6	9	7
Frame “Tensions and confrontation”	3	5	17	6

Figure 4. Popularity of frames in AI-generated texts, by year

Source: compiled by the authors based on ChatGPT data.

Наиболее распространенным стал фрейм «Independent entity», причем показатели его встречаемости в генерируемых текстах оказались ощутимо выше аналогичных показателей других фреймов. Исключением стали показатели 2023 г., где «преимущество» обозначенного фрейма не столь велико. Любопытно, что сгенерированный текст по 2023 г. – это прогнозный текст, когда искусственный интеллект пытается спрогнозировать

ответ на оба используемых текстовых запроса («Taiwan as an independent entity in 2023» и «Taiwan as part of China in 2023»). Если минимальный уровень частоты встречаемости фрейма «Independent entity» приходится на 2023 г., то максимальный уровень – на 2020 г.

Отметим схожую структуру изменений показателей распространенности всех трех выявленных фреймов по разным годам, но с различной амплитудой изменений: в частности, обращает на себя внимание серьезное снижение частоты встречаемости фрейма «Independent entity» в «прогнозном» 2023 г. (по сравнению с 2020 г.), значительное увеличение популярности фрейма «Tensions and confrontation» в 2020 г. (по сравнению с 2015 г.) и практически аналогичное ему снижение популярности указанного фрейма в 2023 г. В целом, подчеркнем, что самым распространенным оказался фрейм «Independent entity» (общее количество знаково-символических единиц анализа равно 86), за ним следуют «Tensions and confrontation» (31 знаково-символическая единица анализа) и «The part of China» (с показателем знаково-символических единиц, равным 26).

Такое отличие и падение количества маркеров в 2023 г. можно объяснить попыткой алгоритмов сделать прогноз, а не использовать исключительно уже созданные массивы текста на данный год. Недостаток текущих данных и несовершенство алгоритмов с точки зрения описания возможных сценариев приводит к изменению характерной картины маркеров и увеличению количества малосодержательных общих фраз в тексте, что подробнее проанализируем в следующей части статьи.

Выводы

Во-первых, по текстовому запросу, напрямую связанному с независимым и суверенным политическим статусом Тайваня, искусственный интеллект генерирует тексты б*о*льшего объема, чем по текстовому запросу, относящемуся к утверждению о его территориальной принадлежности Китайской Народной Республике. Это может свидетельствовать о том, что в англоязычной литературе доминирует рассмотрение вопроса о территориальном статусе Тайваня как независимого, суверенного субъекта. Подчеркнем, в содержании текстовых запросов использованы максимально упрощенные и, на наш взгляд, наиболее релевантные формулировки.

Во-вторых, выявив с учетом используемого контекста слова/выражения-маркеры, поддерживающие один из двух противоположных подходов к пониманию территориального статуса Тайваня, и подсчитав частоту их встречаемости во всех сгенерированных текстах, можно отметить абсолютное преимущество по данному показателю слов/выражений-маркеров, символически отражающих идею о независимом и суверенном политическом статусе Тайваня, а не о его территориальной принадлежности КНР. Разумеется, слова/выражения-маркеры, передающие идею о независимом и суверенном политическом статусе Тайваня, встречаются пре-

имущественно в текстах, сгенерированных по соответствующему запросу. Впрочем, и в текстах, сгенерированных ИИ по обратному смысловому варианту, нередко встречаются подобные семиотические единицы-маркеры. Слова-маркеры могут быть нацелены, прежде всего, на усиление психологической реакции индивида при описании какого-либо процесса или события, способствующего формированию в его сознании определенного ассоциативного ряда. Зачастую, чем больше объем информации, тем выше вероятность использования в ней подобных знаковых единиц (хотя такая закономерность, вероятно, наблюдается далеко не всегда). Подобные слова и выражения можно поделить на типы, причем различные знаковые единицы-маркеры обладают и разным потенциалом влияния на сознание индивида, а в дальнейшем – на его поведение как ответную реакцию на информационное воздействие.

Выявленные слова/выражения-маркеры, отражающие идею о независимом и суверенном политическом статусе Тайваня, можно разделить как минимум на те из них, которые (1) фиксируют непосредственно статус территориальной единицы («state», «country», «nation» и др.) и (2) характеризуют его (и его действия) по отношению к КНР и другим странам («independent», «separate», «international», «diplomatic» и др.). В свою очередь, установленные слова/выражения-маркеры, передающие идею о территориальной принадлежности Тайваня КНР, можно разделить как минимум на те из них, которые также (1) определяют статус территории («part/province of its/China territory», «representative of China» и др.) и (2) описывают действия Китая по отношению к Тайваню («control», «One China Policy» и др.).

Точную причину такого явления определить достаточно сложно, поскольку она может быть следствием как одного фактора, так и нескольких, в том числе намеренных, спровоцированных с целью сформировать конкретную политическую повестку, и ненамеренных, полученных в результате технических ограничений, несовершенства алгоритмов и ограниченности массивов данных.

В-третьих, наибольший объем текста по запросу о независимом и суверенном политическом статусе Тайваня и наибольшее количество соответствующих слов/выражений-маркеров представлено в сгенерированном ИИ тексте 2020 г. Интересно, что наибольший объем текста и по другому запросу, содержащему идею о территориальной принадлежности Тайваня КНР, сгенерирован ИИ также в 2020 г., но соответствующих содержанию запроса знаковых единиц-маркеров в тексте 2020 г. представлено столько же, сколько и в предыдущие годы.

Особенности генерирования ИИ текстов по указанным запросам могут обуславливаться частотой обсуждения указанной темы в интернет-пространстве, что, в свою очередь, вызвано определенным событийным рядом, в частности, победой Демократической прогрессивной партии на парламентских выборах в январе 2020 г., а также уверенной победой ее лидера Цай Инвэнь на президентских выборах того же года,

выступающих за самостоятельность острова[4], визитом на Тайвань спикера чешского Сената Милоша Выстрчила в августе того же года и реакцией на это событие со стороны КНР[5], проведением военных учений КНР в Южно-Китайском море в июле[6] и последующими учениями США в этом же районе[7]. Возможно и наличие других причин подобного генерирования текстов искусственным интеллектом, например, технологического свойства (вызванных спецификой обучения и работы нейронной сети и т. д.) или особенностями текстового запроса (подчеркнем, для поиска информационного массива использовались максимально упрощенные варианты формулировок текстовых запросов).

В-четвертых, говоря о выявленных фреймах и их популярности в проанализированных текстовых данных, отмечаем наибольшую частоту использования именно фрейма «Independent entity», что согласуется с другими результатами, описанными выше. Причем, всплеск популярности обозначенного фрейма также приходится именно на 2020 г. Интересно, что в «прогнозном» тексте, генерируемом ИИ на период 2023 г., указанный фрейм хотя и продолжает оставаться самым распространенным, но частота его использования серьезно уменьшается и не так преобладает над показателями популярности двух других фреймов, как в более ранние периоды. Это может быть вызвано в целом снижением объема текстов, генерируемых ИИ по 2023 г. по сравнению с 2020 г. (хотя по показателю объема слов они находятся на втором месте среди всех четырех временных отрезков), а, возможно, – «стремлением» искусственного интеллекта сделать в будущем генерируемые тексты по данной проблематике более нейтральными по своей тональности, что перекликается и со снижением количества слов/выражений-маркеров, отражающих оба подхода к проблеме политического статуса Тайваня, в обоих текстах, сформированных ИИ по запросам двух типов по 2023 г. Кроме того, причина может заключаться в несовершенстве технической стороны для прогнозирования, что является более сложной математической задачей, нежели классическая обработка данных. Не забываем и о возможном наличии факторов технического свойства, повлиявших на результаты генерирования ИИ «прогнозного» текста 2023 г.

Отдельно отметим, что вторым по популярности стал фрейм «Tensions and confrontation», а не другой фрейм (созвучный содержанию одного из текстовых запросов) – «The part of China». Особую роль в этом сыграла

[4] Китай, уходи! На Тайване взяли власть сепаратисты. У США и Китая появился еще один повод для конфликта // Lenta.ru. 13 января 2020. https://lenta.ru/articles/2020/01/13/taiwan/

[5] В КНР обвинили спикера чешского Сената в нарушении суверенитета республики // Tass.ru. 30 августа 2020. https://tass.ru/mezhdunarodnaya-panorama/9323479

[6] Китай назвал цель маневров в Южно-Китайском море // Газета.ru. 9 июля 2020. https://www.gazeta.ru/army/news/2020/07/09/14650015.shtml?updated

[7] Вслед за КНР: авианосцы США вошли в Южно-Китайское море // Газета.ru. 6 июля 2020. https://www.gazeta.ru/army/2020/07/06/13142863.shtml?updated

степень распространенности обозначенного фрейма именно в тексте 2020 г., а причиной подобной популярности может быть названа специфика событийного ряда, который представлен в одном из пунктов выше.

В целом наблюдается схожая динамика изменения популярности всех выявленных фреймов: рост показателей распространенности с 2010 по 2020 г. со снижением (в двух случаях из трех – со значительным уровнем снижения) к «прогнозному» для ИИ 2023 г.

В ходе выполнения фрейм-анализа была предпринята попытка поиска других фреймов, в частности, «Cooperation», но найти знаковые единицы, отражающие такой фрейм применительно к идее взаимодействия КНР и Тайваня, не удалось. Отчасти подобный фрейм проявляется в представленной в одном из текстов идее о членстве Тайваня во Всемирной торговой организации: «...Taiwan is an active participant in international organizations and events under the name "Chinese Taipei". It is a member of the World Trade Organization under this designation and participates in some other international forums and events, but not as a fully recognized independent state» («Тайвань является активным участником международных организаций и мероприятий под названием «Китайский Тайбэй». Под этим названием он является членом Всемирной торговой организации и участвует в некоторых других международных форумах и мероприятиях, но не как полностью признанное независимое государство»).

Практически никак не представлен в сгенерированном ИИ информационном массиве и фрейм, связанный с гражданской (национальной) идентичностью, не считая отдельных редких упоминаний данной темы («The majority of Taiwanese people identify themselves as Taiwanese rather than Chinese...» («Большинство тайваньцев идентифицируют себя как тайваньцы, а не китайцы...») и т. д.).

Заключение

В данной работе был проанализирован массив текстов, сгенерированных ChatGPT по заданной тематике, и выявлены характерные закономерности, которые могут получить различную интерпретацию и объяснения. В результате исследования обнаружено, что маркеры позиционирования статуса Тайваня как «независимого государства» имеют существенное количественное преимущество по различным временным точкам, нежели маркеры, связанные с позиционированием Тайваня как части Китайской Народной Республики.

Чтобы сделать более определенные выводы о специфике и механизмах работы алгоритмов инструмента, необходимы дальнейшие глубокие и массовые исследования с применением методов математического анализа и моделирования (Petukhov et al., 2018; Petukhov et al., 2020), а также сравнительных исследований, с учетом того, что сам инструмент параллельно будет дорабатываться и изменяться.

Список литературы

Беляков М.В., Максименко О.И. Диалоговые системы: история развития и чат-боты GPT как новая лингвосемиотическая реальность // Военно-гуманитарный альманах: материалы XVII Международной научной конференции по актуальным проблемам языка и коммуникации «Язык. Коммуникация. Перевод». М.: Военный университет Министерства обороны РФ, 2023. С. 477–485.

Боришполец К.П. Методы политических исследований. 2-е изд., испр. и доп. М.: Аспект Пресс, 2010. 230 с.

Вахштайн В.С.[8] Социология повседневности и теория фреймов. СПб.: Изд-во Европейского университета в Санкт-Петербурге, 2011. 334 с.

Гудев П.А., Мишин И.О. Американо-китайские противоречия вокруг Тайваньского пролива (Часть I. Суть терминологического спора) // Юго-Восточная Азия: актуальные проблемы развития. 2022. Т. 3. № 3 (56). С. 60–76.

Дикарев А.Д., Лукин А.В. «Тайваньская нация»: от мифа к реальности? // Сравнительная политика. 2021. Т. 12. № 1. С. 118–133. https://doi.org/10.24411/2221-3279-2021-10009

Жук А.Р. Правовое положение Тайваня // Российское право онлайн. 2023. № 1. С. 75–81. https://doi.org/10.17803/2542-2472.2023.25.1.076-081

Конуров А.И. Современное состояние и перспективы разрешения проблемы политического статуса Тайваня в свете эволюции позиции США // Вестник Московского государственного лингвистического университета. Общественные науки. 2020. № 3 (840). С. 83–97.

Одноконный Г.А., Рогач С.И. Влияние чата GPT на развитие современного общества и бизнеса // Материалы XV Всероссийской научно-практической конференции «Проблемы современного социума глазами молодых исследователей». Волгоград: Университетская книга, 2023. С. 363–364.

Петухов А.Ю., Полевая С.А., Каминченко Д.И. Влияние современных СМИ на политические ценности индивида на примере «RT» и «BBC» // Вестник Томского государственного университета. Философия. Социология. Политология. 2023. № 72. С. 253–271. https://doi.org/10.17223/1998863X/72/22

Пузыня Н.Н. Китайско-американские отношения и проблема Тайваня // Известия Иркутского государственного университета. Серия: Политология. Религиоведение. 2022. Т. 40. С. 56–63. https://doi.org/10.26516/2073-3380.2022.40.56

Субботкин В.Д. Глобальные СМИ и их влияние на мировую политику (по материалам CNN и BBC) / под общ. ред. М.Э. Вильчинской-Бутенко // Труды института бизнес-коммуникаций. СПб.: Санкт-Петербургский государственный университет промышленных технологий и дизайна, 2017. С. 49–54.

Сунь Ц. Дипломатический статус Тайваня (2016–2018 гг.) // Белорусско-российские отношения на современном этапе: состояние и перспективы развития: материалы круглых столов. Минск: БГУ, 2018. С. 54–55.

Фадеева И.А. Особенности взаимоотношений Китая и Тайваня: История вопроса и современные традиции (экономический аспект) // Теория и практика общественного развития. 2023. № 6 (182). С. 192–197.

Шаров К.С. Политический статус-кво КНР и Тайваня в 2010-х: сдвиг в сторону националистической идеологии // Социология. 2018. № 2. С. 220–224.

Baum J., Villasenor J. The politics of AI: ChatGPT and political bias // Brookings.edu. May 8. 2023. https://www.brookings.edu/articles/the-politics-of-ai-chatgpt-and-political-bias/

Chou Ch.-Ch. Contending notions of the cross-strait status quo in Taiwan and across the strait: Impacts on U.S.-Taiwan relations // China Review. 2018. Vol. 18. № 3. P. 121–148.

[8] Внесен Министерством юстиции РФ в реестр СМИ, выполняющих функции иностранного агента.

Lidgren S. Movement mobilization in the age of hashtag activism: examining the challenge of noise, hate, and disengagement in the #MeToo Campaign // Policy & Internet. 2019. Vol. 11. Issue 4. P. 418–438. https://doi.org/10.1002/poi3.212

Petukhov A.Yu., Malkhanov A.O., Sandalov V.M., Petukhov Yu.V. Mathematical modeling ethno-social conflicts with the introduction of the control function // Simulation. 2020. Vol. 96. Issue 3. P. 337–346. https://doi.org/10.1177/0037549719884629

Petukhov A.Yu., Malkhanov A.O., Sandalov V.M., Petukhov Yu.V. Modeling conflict in a social system using diffusion equations // Simulation. 2018. Vol. 94. Issue 12. P. 1–9. https://doi.org/10.1177/0037549718761573

Rauscher N., Schäfer M. Leveling the playing field – Kommunikationsstrategien und Ungleichheitsnarrative des linken Diskurses im Umfeld der Demokratischen Partei der USA // Zeitschrift für Politikwissenschaft. 2022. Vol. 32. Nr. 1. P. 221–249. https://doi.org/10.1007/s41358-021-00275-2

References

Baum, J., & Villasenor, J. (2023). The politics of AI: ChatGPT and political bias. *Brookings.edu*, May 8. https://www.brookings.edu/articles/the-politics-of-ai-chatgpt-and-political-bias/

Belyakov, M.V., & Maksimenko, O.I. (2023). Dialogue systems: history of development and GPT chatbots as a new linguosemiotic reality. *Military and Humanitarian Almanac: Proceedings of the XVII International scientific conference on current problems of language and communication "Language. Communication. Translation"*. Moscow: Military University of the Ministry of Defence of the Russian Federation Publ. P. 477–485. (In Russ)

Borishpolets, K.P. (2012). *Methods of political research*. 2nd ed. Moscow: Aspekt Press. (In Russ.)

Chou, Ch.-Ch. (2018). Contending notions of the cross-strait status quo in Taiwan and across the strait: Impacts on U.S.-Taiwan relations. *China Review*, *18*(3), 121–148.

Dikarev, A.D., & Lukin, A.V. "Taiwan nation": from myth to reality? *Comparative Politics Russia*, *12*(1), 118–133. (In Russ.)

Fadeeva, I.A. (2023). Peculiarities of relations between China and Taiwan: history of the issue and modern traditions (economic aspect). *Theory and Practice of Social Development*, *6*(182), 192–197. (In Russ.)

Gudev, P.A., & Mishin, I.O. (2022). American-Chinese controversies around the Taiwan strait (Part I. Essence of the terminological dispute). *Southeast Asia: Actual Problems of Development, 3*(3), 60–76. (In Russ.)

Konurov, A.I. (2020). Political status of Taiwan problem: current state and prospects of solution in the wake of evolving us stance. *Vestnik of Moscow State Linguistic University (Social sciences)*, *3*(840), 83–97. (In Russ.)

Lidgren, S. (2019). Movement mobilization in the age of hashtag activism: examining the challenge of noise, hate, and disengagement in the #MeToo Campaign. *Policy & Internet*, *11*(4), 418–438. https://doi.org/10.1002/poi3.212

Odnokonny, G.A., & Rogach, S.I. (2023). The influence of ChatGPT on the development of modern society and business. *Proceedings of the XV Russian scientific and practical conference "Problems of modern society through the eyes of young researchers"*. Volgograd: Universitetskaya kniga. P. 363–364. (In Russ.)

Petukhov, A.Yu., Malkhanov, A.O., Sandalov, V.M., & Petukhov, Yu.V. (2018). Modeling conflict in a social system using diffusion equations. *Simulation*, *94*(12), 1–9. https://doi.org/10.1177/0037549718761573

Petukhov, A.Yu., Malkhanov, A.O., Sandalov, V.M., & Petukhov, Yu.V. (2020). Mathematical modeling ethno-social conflicts with the introduction of the control function. *Simulation*, *96*(3), 337–346. https://doi.org/10.1177/0037549719884629

Petukhov, A.Yu., Polevaya, S.A., & Kaminchenko, D.I. (2023). The influence of modern media on the political values of the individual: on the example of RT and BBC. *Tomsk State University Journal of Philosophy, Sociology and Political Science*, *72*, 253–271. (In Russ.) https://doi.org/10.17223/1998863X/72/22

Puzynya, N.N. (2022). Chinese-American relations and problem of Taiwan. *Bulletin of Irkutsk State University ("Political Science and Religion Studies" Series), 40*, 56–63. (In Russ.)

Rauscher, N., & Schäfer, M. (2022). Leveling the playing field – progressive communication strategies and current narratives on inequality in the Democratic Party in the U. S. *Zeitschrift für Politikwissenschaft*, *32*(1), 221–249.

Sharov, K.S. (2018). People's Republic of China and Taiwan political status quo in the 2010s: a shift towards nationalist ideology. *Sociology*, *2*, 220–224. (In Russ.)

Subbotkin, V.D. (2017). Global media and their influence on international policy (based on CNN and BBC). In: Vilchinskaya-Butenko M.E. (ed.) *Scientific papers of the Institute of Business Communications*. St. Petersburg: St. Petersburg State University of Industrial Technologies and Design. P. 49–54. (In Russ.)

Sun' Ts. (2018). Taiwan's diplomatic status (2016–2018). *Belarusian-Russian relations at the present stage: state and development prospects: materials of round table conferences*. Minsk: Belarusian State University. P. 54–55. (In Russ.)

Vakhshtayn, V.S.[9] (2011). *Sociology of everyday life and frame theory*. St. Petersburg: European University at St. Petersburg.

Zhuk, A.R. (2023). The legal status of Taiwan. *Russian Law Online*, *1*, 75–81. (In Russ.)

Сведения об авторах:

Петухов Александр Юрьевич, кандидат политических наук, доцент, заведующий лабораторией математических методов политического анализа и прогнозирования, Московский государственный университет имени им. М.В. Ломоносова, Российская Федерация, 119991, Москва, Ленинские горы, д. 1; заведующий лабораторией когнитивной безопасности, Университет «Неймарк», Российская Федерация, 603057, Нижний Новгород, ул. Нартова, д. 6. ORCID: 0000-0002-7412-5397. E-mail: lectorr@yandex.ru

Каминченко Дмитрий Игоревич, кандидат политических наук, доцент Института международных отношений и мировой истории, Нижегородский государственный университет им. Н.И. Лобачевского, Российская Федерация, 603022, Нижний Новгород, проспект Гагарина, д. 23; старший научный сотрудник лаборатории когнитивной безопасности, Университет «Неймарк», Российская Федерация, 603057, Нижний Новгород, ул. Нартова, д. 6. ORCID: 0000-0002-3193-3423. E-mail: dmitkam@inbox.ru

Bio notes:

Aleksandr Yu. Petukhov, Candidate of Political Sciences, Associate Professor, Head of the Laboratory of Mathematical Methods of Political Analysis and Forecasting, Lomonosov Moscow State University, 1 Leninskie Gory, Moscow, 119991, Russian Federation; Head of the Cognitive Security Lab, Neimark University, 6 Nartova St, Nizhny Novgorod, 603057, Russian Federation. ORCID: 0000-0002-7412-5397. E-mail: lectorr@yandex.ru

Dmitriy I. Kaminchenko, Candidate of Political Sciences, Associate Professor at the Department of Political Science, Institute of International Relations and World History, Lobachevsky State University of Nizhny Novgorod, 23 Gagarina Prospect, Nizhny Novgorod, 603022, Russian Federation; Senior Researcher of the Cognitive Security Lab, Neimark University, 6 Nartova St, Nizhny Novgorod, 603057, Russian Federation. ORCID: 0000-0002-3193-3423. E-mail: dmitkam@inbox.ru

[9] Внесен Министерством юстиции РФ в реестр СМИ, выполняющих функции иностранного агента.