

Марина Николаевна Николаева, к. филол. н.
Московский городской педагогический универ-
ситет, Москва, Россия
э-почта: nikolaevam@mgpu.ru

Marina Nikolaevna Nikolaeva, Cand. of Sc.
(Philology)
Moscow City University, Moscow, Russia
e-mail: nikolaevam@mgpu.ru

Марина Анатольевна Овсянникова, к. филол. н.
Московский городской педагогический универ-
ситет, Москва, Россия
э-почта: ovsyannikovama@mgpu.ru

Marina Anatolyevna Ovsyannikova, Cand.
of Sc. (Philology)
Moscow City University, Moscow, Russia
e-mail: ovsyannikovama@mgpu.ru

ПРИМЕНЕНИЕ ИНСТРУМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ АНАЛИЗЕ ФРАЗЕОЛОГИЧЕСКИХ ЕДИНИЦ С КОМПОНЕНТОМ КОЛОРОНИМОМ

Данная статья представляет результаты применения инструментов искусственного интеллекта при отборе лексического материала, в частности фразеологических единиц с компонентом колоронимом в английском языке.

Ключевые слова: искусственный интеллект; фразеология; идиома; колороним; метод сплошной выборки.

ARTIFICIAL INTELLIGENCE AS A TOOL FOR COLOUR IDIOMS ANALYSIS

This article presents the results of artificial intelligence application for colour idioms analysis in English.

Key words: artificial intelligence; phraseology; idiom; colour idiom; sampling method.

В настоящий момент разработчики платформ искусственного интеллекта (ИИ) позиционируют свои инструменты для поиска и отбора информации как эффективные помощники исследователя, призванные облегчить работу специалистов разных областей, в том числе лингвистов. Цель настоящего исследования – отобрать лексический материал, в частности, фразеологизмы (ФЕ) с компонентом колоронимом, для дальнейшего лингвистического анализа с помощью открытых и доступных в РФ инструментов искусственного интеллекта.

Материалом исследования послужил Оксфордский словарь идиом (The Oxford Dictionary of Idioms, 2004). Данный словарь привлек внимание авторов возможностью работы с неструктурированными данными: объемный текст словаря (352 страницы в формате pdf) представляется идеальным материалом для применения инструментов искусственного интеллекта при отборе и анализе лексического материала. В качестве инструментов ИИ было выбрано 4 платформы: а) SciSpace (URL: <https://typeset.io/>): платформа позиционирует себя как «самая быстрая исследовательская ИИ платформа для студентов и исследователей» и может работать с pdf-файлами; б) Perplexity AI (URL: <https://www.perplexity.ai>): поисковая система и чат-бот на основе искусственного интеллекта; в) Le Chat Mistral (URL: <https://chat.mistral.ai>): чат-бот, основной европейский конкурент широко известного ChatGPT; г) Julius AI (URL: <https://julius.ai/>): инструмент анализа больших объемов данных на основе искусственного интеллекта.

Российские разработки (Giga.Chat и YandexPro) в отборе материала не применялись, поскольку в настоящее время не имеют функции анализа документа в pdf-формате. Китайская нейросеть Deepseek отказалась выполнять запрос ввиду возможного нарушения авторских прав составителей словаря.

Анализ языкового материала проходил в несколько этапов. На первом этапе широко сформулированный промпт (*Retrieve all the colour idioms from the dictionary*

uploaded) не дал релевантных результатов, и авторы были вынуждены сузить свой запрос: *Retrieve all the idioms with the word "green" from the dictionary uploaded*. Ниже представлены ответы платформ ИИ в виде таблицы (см. таблица 1).

Табл. 1 Результаты отбора ФЕ с колоронимом с помощью ИИ платформ

ИИ ПЛАТФОРМА	РЕЗУЛЬТАТЫ
Le Chat Mistral	- the green-eyed monster
SciSpace	- green with envy - the green-eyed monster
Perplexity AI	- give the green light - the grass is always greener (on the other side) - green around the gills
Julius AI	- green around the gills - green-eyed monster - green fingers - green light - green with envy

Для верификации результатов на втором этапе анализа был применен метод сплошной выборки. По результатам традиционного «ручного» анализа в словаре было найдено 7 идиом с колоронимом *green*: *the grass is always greener* – рус. ‘трава всегда зеленее на другой стороне дороги, или у других всегда всё лучше, хотя это может быть не совсем так’; *green about the gills* – рус. ‘выглядеть бледным и больным’; *green light* – рус. ‘разрешение сделать что-л.’; *green with envy* – рус. ‘желать обладать тем, что есть у других, завидовать кому-л.’; *the green-eyed monster* – рус. ‘зависть’; *wigs on the green* – ‘жестокие или неприятные события, ссоры’; *wear the green willow* – рус. ‘горевать об утраченной или безответной любви’.

Как видно из таблицы выше, ни одна из платформ не дала полного ответа, а Julias.Ai выдал идиому, которая отсутствует в словаре (*green fingers* – рус. ‘способность человека выращивать растения’), что свидетельствует о явной галлюцинации. Однако необходимо отметить, что все использованные платформы ИИ без дополнительного запроса сопроводили свои ответы подробными определениями, а SciSpace дал этимологическую справку для идиомы *the green-eyed monster*: «*It is derived from a quote in Shakespeare's play "Othello" where Iago warns about the dangers of jealousy, referring to it as the green-eyed monster*».

Таким образом, быстрота выполнения запроса инструментами искусственного интеллекта не вызывает сомнения, однако результаты должны быть верифицируемы специалистом в области идиоматики. На данном этапе развития доступных инструментов ИИ можно говорить о том, что использование подобного инструментария для анализа неструктурированного текстового материала должно проводиться в сочетании с традиционными «ручными» методами.