

Шевелева Анастасия Алексеевна

студент департамента филологии

Научный исследовательский университет

«Высшая школа экономики»

Санкт-Петербург, Россия

Anastasiia Sheveleva

Student of the Department of Philology

National Research University «Higher School
of Economics»

St. Petersburg, Russia

anastasiia.sheveleva@gmail.com

НЕЙРОСЕТЬ КАК ХУДОЖНИК: ИЗВЛЕЧЕНИЕ ИНФОРМАЦИИ ИЗ ЛИТЕРАТУРНЫХ ТЕКСТОВ ДЛЯ ГЕНЕРАЦИИ ИЛЛЮСТРАЦИЙ

В статье описываются результаты прикладного исследования, направленного на извлечение из литературных текстов дескриптивных фрагментов, относящихся к героям и ключевым событиям произведения. Целью является разработка алгоритма, который позволит автоматически создавать текстовые запросы (промпты) для генерации иллюстраций к выбранному произведению с помощью нейросетевых технологий. В качестве материалов, на основе которых осуществляется апробация экспериментального дизайна, используются популярные сказки для детей (Я. и В. Гримм «Рапунцель», Ш. Перро «Золушка, или хрустальная туфелька»). Язык написания программы – Python; язык материалов исследования – русский. Для реализации программного решения используются методы извлечения именованных сущностей, алгоритмы экстрактивной суммаризации, а также языковая модель YandexGPT. Для создания иллюстраций и оценки качества

итоговых промптов в рамках поставленной задачи используется генеративная модель YandexART. Полученные результаты имеют практическую значимость и могут быть полезны в вопросе оптимизации процесса генерации иллюстраций при помощи искусственного интеллекта в области креативных индустрий и медиа.

Ключевые слова: *обработка естественного языка; извлечение описания персонажа; промпт-инжиниринг; генерация иллюстраций; нейросеть YandexArt; искусственный интеллект.*

NEURAL NETWORK AS AN ARTIST: INFORMATION EXTRACTION FROM LITERARY TEXTS FOR GENERATING ILLUSTRATIONS

The article describes the results of applied research aimed at extracting descriptive fragments from literary texts relating to the characters and key events of the work. The aim of the work is to develop an algorithm that will allow to automatically create text queries (prompts) to generate illustrations for the selected work with the help of neural network technologies. Popular fairy tales for children (J. and V. Grimm's «Rapunzel», Sh. Perrault's «Cinderella, or the Crystal Slipper») are used as materials on the basis of which the experimental design is tested. The language of programme writing is Python; the language of the research materials is Russian. To implement the software solution the methods of named entity extraction, extractive summarisation algorithms and the YandexGPT language model are used. The generative model YandexART is used to create illustrations and assess the quality of the final prompts within the framework of the task. The obtained results have practical significance and can be useful in optimising the process of generating illustrations with the help of artificial intelligence in the creative industries and media.

Key words: *natural language processing; character description extraction; prompt engineering; illustration generation; YandexArt neural network; artificial intelligence.*

Системы искусственного интеллекта в последнее время получили широкое применение среди специалистов в области медиапроизводства [1]. В частности, отмечается все большее распространение использования нейросетей для генерации изображений. Однако активное внедрение этих технологий затрудняется спецификой работы с ними. Их использование основывается на преобразовании пользовательского промпта (*prompt*) в изображение. Под промптом понимается определенным образом сформулированный запрос, который предоставляется нейросетевой модели перед генерацией данных [2]. Точность и понятность этого запроса напрямую влияют на качество получаемого результата, а потому с развитием нейросетевых технологий появилась новая дисциплина – промпт-инжиниринг (*prompt engineering*).

Промпт-инжиниринг представляет собой процесс разработки и оптимизации промптов для повышения эффективности работы программы [3]. Разные задачи требуют акцентирования внимания на разных деталях, и промпт-инжиниринг позволяет понять, как именно нужно сформулировать программе команду во избежание получения неприменимой информации. Промпты для задачи генерации изображений, именуемые в рамках данной работы «иллюстративные промпты», отличаются от промптов, конструируемых для больших языковых моделей (*large language models*, или *LLM*) –

моделей машинного обучения, умеющих обрабатывать и генерировать текст. Основное отличие состоит в том, что LLM способны считывать и выполнять большой перечень задач, тогда как иллюстративные нейросети ограничены только одной областью, что обуславливает специфику правил, которым необходимо следовать при написании запросов.

Чтобы проиллюстрировать эти ограничения, можно обратиться к инструкциям, выдаваемым конкретными сервисами. Так, для YandexART – «картиночной» генеративной модели от «Яндекса» – составлен список советов, который выглядит следующим образом:

- 1) «Используйте яркие образы...»;
- 2) «Указывайте больше деталей для объектов, такие как цвет, форма, расположение и другие»;
- 3) «Указывайте стили, если хотите получить изображение определенного вида»;
- 4) «Заменяйте редкие и сложные концепции на их описание» [4].

Ниже представлен пример промпта из раздела «Иллюстрации для детей», составленный в соответствии с предложенными рекомендациями, и сгенерированное на его основе изображение (рис.1).



Рис. 1. Пример иллюстрации из библиотеки промптов YandexART

«Мультяшная лисичка сидит на подоконнике, смотрит в окно, рядом чашка чая, теплый вечер за окном, закат, уют, ламповая атмосфера, мягкие яркие цвета, простые текстуры, плоская иллюстрация из детской книжки, рисование водными маркерами» [4].

При этом можно также соблюдать общие рекомендации, применимые как к промптам для LLM, так и к иллюстративным промптам. Главное из них – конкретность и точность передаваемых инструкций. «Чем более подробным и детальным будет промпт, тем лучше будут и результаты» [3]. Однако стоит учитывать, что если текст будет слишком длинным, то модель не сможет обработать его полностью и просто упустит часть информации.

Существуют научные работы, в которых анализируются проблемы, связанные с промпт-инжинирингом [5]. В некоторых исследованиях изучаются возможности автоматической оптимизации процесса создания промптов [6], осуществляются попытки разработки универсальной модели генерации промптов без вовлечения человека [7]. Однако иллюстративные промпты в таких работах не затрагиваются, вследствие чего возникает вопрос об уровне сложности этой задачи и о возможности ее реализации в целом. При этом автоматизация генерации иллюстративных промптов несет высокую прак-

тическую значимость, поскольку позволит оптимизировать процесс генерации иллюстраций в промышленных целях – например, для создания и выпуска книг или учебных пособий.

В этой работе осуществлена попытка создания алгоритма, позволяющего осуществить частичную автоматизацию составления иллюстративных промптов. Полученный алгоритм направлен на извлечение из литературных текстов дескриптивных фрагментов, относящихся к героям и ключевым событиям. Цель реализации этой программы – получение данных, на основе которых можно вручную написать текстовые запросы (промпты) для генерации иллюстраций к выбранному произведению с помощью нейросетевых технологий.

Для достижения поставленной цели решаются следующие задачи:

- подбор требуемых методов извлечения и обработки информации;
- создание вручную иллюстративных промптов на основе полученных данных;
- апробация промптов при помощи нейросетевых технологий;
- оценка полученных иллюстраций.

В результате выполнения поставленных задач на данном этапе получена программа, способная автоматически извлекать необходимые для создания иллюстративного промпта дескриптивные фрагменты. Осуществление подобной программы является значимым этапом на пути к реализации полностью автоматизированного алгоритма генерации иллюстративных промптов, потому как позволяет решить большой перечень необходимых для этого задач. Кроме того, полученный алгоритм может быть полезен для извлечения значимой информации из текста с целью его быстрого изучения, составления краткого описания.

В качестве материалов будут использоваться популярные сказки (Я. и В. Гримм «Рапунцель» [8], Ш. Перро «Золушка, или хрустальная туфелька» [9]). Выбор объясняется следующими соображениями: образы из этих произведений знакомы многим, что облегчит оценку качества работы программы. Кроме того, мы опираемся на идею о том, что детским текстам свойственна упрощенная речь и меньшее количество метафорических описаний, которые могут значительно усложнить процесс анализа произведения. Это является важным аспектом на начальных этапах работы по созданию модели генерации промптов. Язык написания модели – Python; язык материалов исследования – русский; программа, подобранная для проверки итоговых промптов и создания иллюстраций, – YandexART [10].

Этапы реализации алгоритма

Первый этап. Суммаризация

Рассмотрим поэтапную работу алгоритма. Поскольку каждая из иллюстраций должна демонстрировать конкретный эпизод в книге, первая задача, которую должен выполнять алгоритм, – определение предложений, подлежащих иллюстрированию. Для реализации этой задачи можно исполь-

зовать метод экстрактивной суммаризации. С его помощью из текста будут извлечены наиболее информационно значимые предложения, из которых позднее будут составлены промпты.

Сначала определим количество эпизодов – изменяемый параметр, который можно регулировать и который зависит от желаемого пользователем количества иллюстраций к произведению. Далее осуществим экстрактивную суммаризацию при помощи библиотеки `sumy` и графового метода `TextRank`, который определяет наиболее значимые предложения посредством построения взвешенного графа, где вершинами являются предложения, а ребрами – их сходство [11, с. 150]. В зависимости от схожести отдельного предложения с другими определяется его вес, самые «тяжелые» предложения – самые информационно важные. При этом, настраивая объем выдачи, важно сделать запас по количеству предложений, поскольку некоторые варианты из суммаризированной выборки нам могут не подойти (в них будет отсутствовать главный герой, что будет пояснено позднее).

В т о р о й э т а п. Определение главных героев

Далее следует учесть, что в иллюстрациях к книгам важно, чтобы изображение не противоречило содержанию текста. Сложности в данном вопросе в основном вызывают персонажи, поскольку случайным образом созданный искусственным интеллектом герой может значительно отличаться по внешности от того, что давалось в тексте за рамками иллюстрируемого предложения. Соответственно, программа должна уметь «узнавать» персонажей в выбранном предложении и «определять» их внешность. Для первой из указанных задач мы используем извлечение именованных сущностей. Его мы осуществим при помощи библиотеки `Natasha`, содержащей модули для работы с русскоязычными текстами, которые позволяют осуществить извлечение именованных сущностей, а также морфологический и синтаксический анализы [12]. Используя необходимые модули, проходимся по каждому слову в тексте, присваиваем ему тег `PER`, если это человек, и выводим лемму в заранее созданный пустой список. После этого подсчитываем упоминание каждой леммы в списке и выводим нужное количество (его можно отрегулировать в зависимости от пожеланий) наиболее частотных единиц – это будут наши главные герои.

Далее подготавливаем итоговый список предложений для создания промптов: исключаем предложения, в которых не встречаются главные герои. Для этого проверяем наличие в предложении из суммаризированного текста имени главного героя. В завершение этого этапа подсчитываем количество полученных предложений и сравним его с желаемым числом иллюстраций, которое мы определяли ранее. Если получившееся значение больше желаемого, исключим ненужные предложения (все, что идет после выбранного нами числа предложений), если меньше – настроим суммаризацию на большую выдачу, если числа совпадают – оставим предложения без изменений.

Т р е т и й э т а п. Извлечение описания внешности персонажей

Для определения внешности персонажей необходимо извлечь из текста всю описательную информацию о них. Эти данные мы будем получать при помощи нейросетей. Для этого обратимся к технологиям YandexGPT и API, позволяющим отсылать запросы в систему YandexGPT. Важно заметить, что необходимо заранее подготовить запрос к системе (далее он не подвергается изменениям без необходимости), который сформулирует конкретную задачу для YandexGPT. Итоговый промпт выглядит следующим образом:

«Ты робот, который занимается извлечением описаний внешности персонажей из текста. Тебе дается сначала имя персонажа, потом текст. Ты должен найти все описание внешности указанного персонажа в тексте. Выводи только описание».

Четвертый этап. Формирование иллюстративных промптов

В завершение, объединяя описание внешности нужного персонажа с конкретным предложением, мы вручную (учитывая все рекомендации, описанные во введении) создадим промпты, на основе которых будут генерироваться иллюстрации. Необходимость осуществления этого этапа вручную имеет следующие причины:

- наличие местоимений в итоговой выборке предложений (автоматическое определение того, к какому субъекту относится местоимение, невозможно без использования методов машинного обучения, а они, в свою очередь, требуют наличие обучающей выборки);

- наличие лишней информации от системы в описании, полученном при помощи YandexGPT. Под лишней информацией подразумеваются общие фразы, при помощи которых нейросеть формулирует связанный ответ. Например, «В тексте <...> приводится следующее описание внешности главной героини:<...>». Ее извлечение аналогично задаче извлечения описания из текста в целом, поскольку формулировки выдачи YandexGPT непредсказуемы. Соответственно, единственный вариант обработки описания без создания дополнительных словарей – вручную.

Апробация предлагаемого решения

1. Я. и В. Гримм «Рапунцель»

Проведем апробацию модели на примере сказки братьев Гримм «Рапунцель».

При помощи алгоритма успешно извлекается имя главной героини – Рапунцель. В результате мы получаем следующую выборку предложений, которые будут использоваться для генерации иллюстраций:

1) *«Тогда у Рапунцель и страх прошел, и когда он ее спросил, желает ли она его взять себе в мужья, а она видела, что он и молод, и прекрасен, то подумала: «Ему я буду милее, чем старой Гошэль» (так звали волшебницу);*

2) *В гневе своем ухватила она чудные волосы Рапунцель, обмотала их два раза около левой руки, а в правую взяла ножницы, и – раз, ее волосы были обрезаны, и чудные косы пали к ногам волшебницы;*

3) *Тогда он вдруг услышал голос, показавшийся ему знакомым; он пошел прямо на этот голос, и, когда приблизился, Рапунцель узнала его и со слезами бросилась ему на шею.*

Извлеченное при помощи YandexGPT описание выглядит так (ниже приведен пример полного ответа программы):

В тексте сказки братьев Гримм «Рапунцель» приводится следующее описание внешности главной героини:

Рапунцель была «прелестнейшая девочка».

Также в тексте есть упоминание о волосах Рапунцель, которые были «чудными, тонкими, как золотая пряжа».

Результат выглядит достоверно, однако «прелестнейшая девочка» может подразумевать описание как внешности, так и характера. Более того, с точки зрения иллюстрации оценочные описания внешности неэффективны, так как они не дают точного понимания задачи. Также требуется отметить, что программа не засчитала еще одну немаловажную деталь: «... тогда ее волосы золотистой волной падали на двадцать локтей...» [8]. В этом отрывке даются две самые значимые характеристики волос героини: длина и цвет.

Итак, мы имеем три предложения и неполное описание волос героини. На основе общих рекомендаций и информации, которую можно извлечь из итоговых данных, сформулируем три промпта. Варианты промптов для соответствующих иллюстраций на основе извлеченных данных выглядят так:

1) *«Мужчина делает предложение девушке. У девушки тонкие, как золотая пряжа, волосы. Мужчина улыбается. Девушка улыбается. Теплые цвета, мультяшный рисунок»;*

2) *«Женщина с нахмуренными бровями отрезает волосы плачущей девушке, серые оттенки, мрачная картина, мультяшный рисунок»;*

3) *«Мужчина и девушка обнимаются. Девушка плачет. У девушки тонкие, как золотая пряжа, волосы. Мультяшный рисунок».*



Рис. 2. «Мужчина делает предложение девушке. У девушки тонкие, как золотая пряжа, волосы. Мужчина улыбается. Девушка улыбается. Теплые цвета, мультяшный рисунок»

Полученные на основе этих промптов иллюстрации приведены ниже (рис. 2, рис. 3, рис. 4). Учитывая исходные данные, две из трех иллюстраций можно назвать в некоторой степени успешными: эмоция и общее действие передано верно, присутствует лишь одна грубая фактическая ошибка – цвет волос и их длина, то есть главные отличительные черты героини изображены неверно (из-за отсутствия информации о них в промптах). Второй же запрос не позволяет распознать системе двух отдельных участников событий. Возможно, проблема в насильственном характере действий, поскольку более сокращенный вариант промпта («женщина отрезает волосы девушке») программа блокирует, отказываясь его отображать.



Рис. 3. «Женщина с нахмуренными бровями отрезает волосы плачущей девушке, серые оттенки, мрачная картина, мультяшный рисунок»



Рис. 4. «Мужчина и девушка обнимаются. Девушка плачет. У девушки тонкие, как золотая пряжа, волосы. Мультяшный рисунок»

2. Ш. Перро «Золушка, или хрустальная туфелька»

Теперь рассмотрим работу модели на примере сказки Ш. Перро «Золушка, или хрустальная туфелька».

Имя главной героини вновь извлекается безошибочно – Золушка. Результат суммаризации выглядит так:

1) *Даже король, который был очень стар и больше дремал, чем смотрел по сторонам, и тот открыл глаза, поглядел на Золушку и сказал королеве вполголоса, что давно уже не видел такой обворожительной особы;*

2) *А Золушка в это время разыскала своих сестер, под села к ним и, сказав каждой несколько приятных слов, угостила их апельсинами и лимонами, которые поднес ей сам принц;*

3) *На другой вечер сестры опять отправились во дворец – и Золушка тоже;*

4) *Он усадил Золушку в кресло и, надев хрустальную туфельку на ее маленькую ножку, сразу увидел, что больше примерять ему не придется: башишочок был точь-в-точь по ножке, а ножка – по башишочку.*

Мы не прибегаем к использованию метрик, оценивающих результат суммаризации, однако можно предположить, что в этом случае некоторые из предложений не особо информационно насыщенные.

Например, предложение (3) не несет сюжетно значимой информации.

При помощи YandexGPT была получена выдача, содержащая большой пласт информации, который не относится конкретно к внешности героини. Из всего ответа, предоставленного системой, к данным, которые соответствуют нашему запросу, относится следующий отрывок:

<...> у нее «тонкий, благородный стан», «бледное, доброе лицо». Девушка носит старое, испачканное золой платье <...>. У Золушки прекрасные каштановые волосы, а глаза «сияют, как звёзды».

Промпты, доработанные вручную, имеют следующие формулировки:

1) *«Старый король смотрит на девушку с каштановыми волосами и улыбается. Девушка в пышном нарядном бальном платье. Дворец, бал. Светлые цвета, мультяшный рисунок»;*

2) «Худая девушка с бледной кожей и каштановыми волосами улыбается и протягивает в руке апельсины и лимоны двум девушкам. Мультяшный рисунок»;

3) «Три девушки стоят во дворце. Одна из девушек худая с бледной кожей и каштановыми волосами. Девушки улыбаются. Светлые цвета, мультяшный рисунок»;

4) «Девушка в рваном старом платье сидит в кресле и мерит одну хрустальную туфельку на правую ногу. Девушка худая, у нее каштановые волосы. Мультяшный рисунок».

Иллюстрации, созданные на основе этих промптов представлены ниже (рис. 5, рис. 6, рис. 7, рис. 8).



Рис. 5. «Старый король смотрит на девушку с каштановыми волосами и улыбается. Девушка в пышном нарядном балльном платье. Дворец, балл. Светлые цвета, мультяшный рисунок»



Рис. 6.
«Худая девушка с бледной кожей и каштановыми волосами улыбается и протягивает в руке апельсины и лимоны двум девушкам. Мультяшный рисунок»

Можно отметить, что получившаяся выборка менее удачна, чем изображения, извлеченные к «Рапунцель». Возможно, это объясняется бóльшим количеством описательных деталей, которые не были определены в тексте. Например, обстановка и время, в которое происходит действие: уточнение этой информации повысило бы качество итоговых иллюстраций. Также у системы возникают проблемы с пространственными отношениями между объектами, даже при дополнительном уточнении данных. Таким образом, только одну из генерированных к «Золушке» иллюстраций можно назвать в некоторой степени удачной: изображение к третьему промпту, однако даже оно не в полной мере корректно, потому как сестры на нем имеют одинаковую внешность, что противоречит содержанию сказки.



Рис. 7. Три девушки стоят во дворце. Одна из девушек худая с бледной кожей и каштановыми волосами. Девушки улыбаются. Светлые цвета, мультяшный рисунок»



Рис. 8. «Девушка в рваном старом платье сидит в кресле и мерит одну хрустальную туфельку на правую ногу. Девушка худая, у нее каштановые волосы. Мультяшный рисунок»

3. Оценка результатов

Полученный алгоритм дает возможность вычленивать из художественного текста наиболее сюжетно значимые предложения, определить главных героев в произведении и извлечь описание их внешности. Иными словами, программа реализует все поставленные изначально задачи.

Тем не менее алгоритм не способен выполнять их безошибочно, поскольку на разных этапах реализации программы возникают проблемы. Так, мы сталкиваемся с отсутствием контроля выдачи на этапе извлечения описания внешности персонажей. Это негативно сказывается на точности и полноте получаемых данных и напрямую влияет на качество итоговой иллюстрации и ее соответствие тексту.

В этой работе, опираясь на существующие методы автоматического анализа текста, мы предприняли попытку создания программы, способной извлекать информацию из художественного произведения с целью генерации иллюстративных промптов. Для этого в первую очередь были изучены необходимые практики, после чего был разработан алгоритм, объединяющий различные методы (суммаризацию, NER, YandexGPT). Далее полученная программа была испытана на сказке Ш. Перро «Золушка, или хрустальная туфелька» и сказке братьев Grimm «Рапунцель», что позволило выявить недостатки модели. К ним относятся отсутствие контроля выдачи YandexGPT и ее неточность. Эти аспекты значительно понизили итоговую эффективность алгоритма, но не сделали его абсолютно нерезультативным, поскольку программа все еще способна воплощать большую часть требуемого анализа.

Полученные результаты имеют практическую значимость и могут быть полезны в вопросе оптимизации процесса генерации иллюстраций при помощи искусственного интеллекта в области креативных индустрий

и медиа. Возможность доработки и усовершенствования модели, а также ее полная автоматизация, не реализуемая на текущем этапе исследования, является задачей будущих работ.

ЛИТЕРАТУРА

1. Рубцова Н. В. Нейросети в медиа: возможности, проблемы, перспективы для будущих медиаспециалистов // Вопросы теории и практики журналистики. Иркутск : Научный журнал Байкальского государственного университета, 2024. Т. 13, № 1. С. 156–171.
2. Brex's Prompt Engineering Guide [Electronic resource]. URL: <https://github.com/brexhq/prompt-engineering?tab=readme-ov-file#what-is-a-prompt> (accessed: 05.06.2024).
3. Руководство по промпт-инжинирингу [Электронный ресурс]. URL: <https://www.promptingguide.ai/ru> (дата обращения: 05.06.2024).
4. Библиотека промптов YandexART [Электронный ресурс] // Yandex Cloud. URL: <https://yandex.cloud/ru/docs/foundation-models/prompts/yandexart/> (дата обращения: 05.06.2024).
5. Prompt Problems: A New Programming Exercise for the Generative AI Era / T. Amarouche, B. A. Becker, P. Denny, J. Leinone, A. Luxton-Reilly, B. N. Reeves, J. Prather // SIGCSE 2024, March 20–23. Portland, 2024. P. 296–302.
6. PromptMaker: Prompt-based Prototyping with Large Language Models / E. Jiang, K. Olson, E. Toh, A. Molina, A. Donsbach, M. Terry, C. J Cai // HI: Conference on Human Factors in Computing System. New Orleans, 2022. P. 1–8.
7. Juneja G., Sharma A. A Universal Prompt Generator for Large Language Models // R0-FoMo: Workshop on Robustness of Few-shot and Zero-shot Learning in Foundation Models at NeurIPS 2023. New Orleans, 2023.
8. Гримм В., Гримм Я. Рапунцель [Электронный ресурс]. URL: <https://litrka.ru/bratya-grimm-rapunzel/> (дата обращения: 05.06.2024).
9. Перро Ш. Золушка, или хрустальная туфелька [Электронный ресурс]. URL: https://nukadeti.ru/skazki/zolushka_ili_khrustalnaya_tufelka (дата обращения: 05.06.2024).
10. YandexART [Электронный ресурс] // Yandex Cloud. URL: <https://console.yandex.cloud/folders/b1g0362jb3qbr0caloes/foundation-models/yandexart> (дата обращения: 05.06.2024).
11. Перелетов К. С. Обзор методов суммаризации текстов и области их применения // Высшая школа: научные исследования : материалы Межвузовского международного конгресса. М., 2021. С. 147–156.
12. Проект Natasha – набор Python-библиотек для обработки текстов на естественном русском языке [Электронный ресурс]. URL: <https://natasha.github.io/> (дата обращения: 05.06.2024).