

Оригинальная статья / Original article

<https://doi.org/10.21869/2223-1536-2025-15-4-35-49>



УДК 004.77

Семантически-инвариантное кондиционирование диффузионных моделей: унифицированный фреймворк для кросс-модельного позитивного промптинга

А. А. Зоткина¹, А. И. Мартышкин¹✉, А. А. Павлов¹, А. В. Ткаченко¹

¹ Пензенский государственный технологический университет
пр. Байдукова / ул. Гагарина, д. 1а/11, г. Пенза 440039, Российская Федерация

✉ e-mail: alexey314@yandex.ru

Резюме

Цель исследования – разработка универсальной методологии позитивного промпт-инжиниринга для генерации изображений диффузионными моделями, основанной на глубоком лингво-семантическом анализе взаимодействия «человек – искусственный интеллект» и выявлении кросс-модельных инвариантов.

Методы. В рамках данного исследования применялся междисциплинарный научный подход, объединяющий методы когнитивного анализа и эмпирической верификации.

Результаты. Результаты исследования подтвердили высокую эффективность предложенной универсальной методологии позитивного промпт-инжиниринга, которая значительно повысила качество генерации изображений диффузионными моделями. Экспериментальные данные показали, что промпты, сформированные по разработанной структуре и стратегиям лексической оптимизации, обеспечивают лучшее соответствие заданным характеристикам и более стабильные результаты across различных моделей, при этом статистически значимо превышая качество неструктурированных промптов ($p < 0,01$). Использование многоуровневой системы компонентов и имплицитных методов контроля позволило снизить вариативность нежелательных артефактов, повысить точность передачи визуальных характеристик и упростить процесс формирования промптов, делая его более предсказуемым, воспроизводимым и универсальным для различных платформ. В целом внедрение этой методологии способствует улучшению взаимодействия человека с искусственный интеллект, повышению стабильности и качества визуальных результатов, а также облегчает адаптацию промптов под разные модели и задачи.

Заключение. Проведенное исследование подтвердило эффективность предложенной универсальной методологии позитивного промпт-инжиниринга для генерации изображений диффузионными моделями. Внедрение структурированного подхода и стратегий лексической оптимизации позволяет значительно повысить качество, стабильность и предсказуемость результатов, а также снизить количество нежелательных артефактов. Такой подход способствует более управляемому и универсальному взаимодействию человека с искусственный интеллект, облегчая создание высококачественных изображений в различных моделях и условиях. В дальнейшем использование разработанной методологии может стать основой для повышения эффективности автоматизированных систем генерации визуального контента и расширения их практических возможностей.

Ключевые слова: диффузионная модель; генерация изображений; искусственный интеллект; промпт-инжиниринг; обработка естественного языка; семантический анализ.

Конфликт интересов: Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

© Зоткина А. А., Мартышкин А. И., Павлов А. А., Ткаченко А. В., 2025

Для цитирования: Семантически-инвариантное кондиционирование диффузионных моделей: унифицированный фреймворк для кросс-модельного позитивного промптинга / А. А. Зоткина, А. И. Мартышкин, А. А. Павлов, А. В. Ткаченко // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. 2025. Т. 15, № 4. С. 35–49. <https://doi.org/10.21869/2223-1536-2025-15-4-35-49>

Поступила в редакцию 11.10.2025

Подписана в печать 10.11.2025

Опубликована 26.12.2025

Semantically invariant conditioning of diffusion models: a unified framework for cross-model positive prompting

Alena A. Zotkina¹, Alexey I. Martyshkin¹ ,
Akim A. Pavlov¹, Alexandra V. Tkachenko¹

¹ Penza State Technological University
1a/11 Baidukova Pass. / Gagarina Str., Penza 440039, Russian Federation

 e-mail: alexey314@yandex.ru

Abstract

The purpose of research is development of a universal methodology of positive industrial engineering for image generation by diffusion models based on a deep linguistic and semantic analysis of Human-AI interaction and identification of cross-model invariants.

Methods. Within the framework of this study, an interdisciplinary scientific approach was applied, combining methods of cognitive analysis and empirical verification.

Results. The results of the study confirmed the high efficiency of the proposed universal methodology of positive industrial engineering, which significantly improved the quality of image generation by diffusion models. Experimental data have shown that prompts formed according to the developed structure and lexical optimization strategies provide better compliance with the specified characteristics and more stable results across different models, while statistically significantly exceeding the quality of unstructured prompts ($p < 0,01$). The use of a multi-level system of components and implicit control methods has made it possible to reduce the variability of unwanted artifacts, increase the accuracy of visual characteristics, and simplify the process of creating designs, making it more predictable, reproducible, and universal for various platforms. In general, the implementation of this methodology improves human interaction with AI, increases the stability and quality of visual results, and facilitates the adaptation of products to different models and tasks.

Conclusion. The conducted research has confirmed the effectiveness of the proposed universal methodology of positive industrial engineering for image generation by diffusion models. The introduction of a structured approach and lexical optimization strategies can significantly improve the quality, stability and predictability of results, as well as reduce the number of unwanted artifacts. This approach promotes more manageable and universal human-AI interaction, making it easier to create high-quality images in various models and conditions. In the future, the use of the developed methodology can become the basis for improving the efficiency of automated visual content generation systems and expanding their practical capabilities.

Keywords: diffusion model; image generation; artificial intelligence; industrial engineering; natural language processing; semantic analysis.

Conflict of interest: The Authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

Received 11.10.2025

Accepted 10.10.2025

Published 26.12.2025

Введение

Эпоха четвертой промышленной революции ознаменована беспрецедентным прогрессом в области искусственного интеллекта (ИИ), кульминацией которого стало появление высокоеффективных генеративных моделей [1]. Среди них особое место занимают диффузионные модели, продемонстрировавшие революционные возможности в синтезе фотorealистичных и стилистически разнообразных изображений [2]. Эти модели, основанные на сложных стохастических процессах последовательного добавления и удаления шума, стали де-факто стандартом в индустрии генеративного искусства, компьютерной графики и мультимедийных технологий [3].

Центральным элементом взаимодействия пользователя с данными системами является промпт (prompt) – текстовая инструкция, вербализующая творческий замысел и направляющая сложный процесс генерации к желаемому результату [4]. Качество, точность и семантическая насыщенность промпта напрямую детерминируют финальное визуальное воплощение, превращая процесс его составления – промпт-инжиниринг – в критически важную компетенцию [5].

Эволюция промпт-инжиниринга прошла путь от элементарных описательных фраз до изощренных, многокомпонентных конструкций, включающих управление стилем, композицией, освещением, ракурсом и даже эмоциональной тональностью изображения [6]. В связи с этим возрастает значение промпт-инжиниринга как ключевой компетенции, позволяющей точно и эффективно управлять процессом генерации изображений, обеспечивая создание более точных, стилистически насыщенных и эмоционально выразительных визуальных решений [7].

Однако стремительное развитие и диверсификация диффузионных моделей (Imagen 3-002 [8], DALL-E 3 [9], Flux 1.1 ultra [10], Midjourney V7 [11], Ideogram 3.0 [12], Kandinsky 3.0 [13], Recraft V3 [14] и др.) породили существенную проблему: фрагментацию методологий промпtingа. Каждая модель обладает уникальной архитектурой, обучалась на специфических датасетах и, как следствие, демонстрирует идиосинкритические особенности в интерпретации текстовых инструкций [15]. Возникновение модель-специфичных команд, систем весов, параметров и предпочтительных структур промптов создает

значительные барьеры для пользователей, затрудняя перенос навыков между платформами и требуя непрерывного переобучения [16].

Актуальность настоящего исследования обусловлена острой необходимостью разработки универсальной методологии промпт-инжиниринга, способной обеспечить стабильно высокое качество и предсказуемость генерации изображений на широком спектре современных диффузионных моделей, минимизируя при этом зависимость от специфических для каждой модели настроек. Особую сложность представляет задача формулирования такой методологии в рамках строгих ограничений: использование исключительно позитивных промптов (описывающих желаемые элементы, а не исключающих нежелательные) на английском языке, без применения негативных инструкций, числовых весов и дополнительных параметров модели. Данные ограничения, хотя и усложняют задачу, нацелены на создание фундаментально чистого, семантически ориентированного подхода, не зависящего от технических «костылей», специфичных для отдельных реализаций.

Целью данной научной работы является разработка, теоретическое обоснование и представление универсальной методологии позитивного промпт-инжиниринга, основанной на глубоком лингво-семантическом анализе взаимодействия «человек – ИИ» и выявлении кросс-модельных инвариантов в

механизмах интерпретации промптов диффузионными моделями.

Материалы и методы

Основываясь на теоретических принципах, анализе моделей и выявленных инвариантах [17], предлагается универсальная методология позитивного промпт-инжиниринга.

Лингво-семантические детерминанты эффективного промпта:

1. Синтаксическая структура, которая определяется порядком слов, важностью расположения в начале, грамматической корректностью и логической сегментацией.

2. Семантическая точность, проявляющаяся в конкретности лексики, семантической непротиворечивости и полноте описания.

3. Прагматическая адекватность, предполагающая явное указание контекста и интенции, использование общепринятой терминологии (например, искусство, фото), а также описание визуальных характеристик.

Предлагается иерархическая четырехкомпонентная структура, организованная по принципу убывания семантической важности:

1. [Component 1: Subject Definition] – определение базового концепта / объекта. Определение базового концепта / объекта. Краткое, точное описание основного субъекта, ключевые атрибуты. Приоритет для внимания модели, фундаментально для генерации [18].

2. [Component 2: Style & Medium Specification] – стилистическая и медиумная спецификация, определяет визуальный язык, эстетику. В содержание входит указание стиля (например, "in the style of...", "Art Nouveau illustration"), медиума ("oil painting", "photorealistic photograph") или описание характеристик ("clean vector art"). Является важнейшим аспектом управления генерацией.

3. [Component 3: Contextual Enrichment & Composition] – контекстуальное обогащение и композиция, добавляет детали окружения, фона, композицию. В содержание входит описание фона (например "serene forest background"), дополнительных объектов, композиции ("wide-angle shot", "close-up view", "rule of thirds composition"). Обоснование этого компонента заключается в определении пространственной организации и нарратива, что помогает создать более насыщенное и выразительное изображение.

4. [Component 4: Lighting, Atmosphere & Quality Detailing] – детализация освещения, атмосферы и качества. Он предназначен для уточнения освещения, настроения и визуальных характеристик изображения. В содержание входит описание освещения ("soft natural daylight", "dramatic chiaroscuro lighting"), атмосферы ("moody atmosphere", "vibrant feeling"), качественных характеристик ("highly detailed texture", "sharp focus", "4K resolution", "masterpiece"). Обоснование этого компонента заключается в том, что освещение и атмосфера

кардинально влияют на восприятие изображения, а указание модификаторов качества служит сигналом к более тщательной проработке деталей.

Компоненты разделяются запятыми, формируя единое предложение на английском языке.

Принципы лексической оптимизации для кросс-модельной совместимости включают использование универсального лексикона, предполагающего применение терминов с устоявшимся визуальным значением в области искусства, фотографии и дизайна, таких как Impressionism, Surrealism, Portrait photography, Macro photography, Rule of thirds, Golden hour, Soft light, Hard light, Highly detailed, Sharp focus и др.; описание визуальных характеристик, дополнение или замена названий стилей и эффектов их конкретными описаниями; избегание модель-специфичного жаргона для повышения универсальности; конкретизацию абстрактных дескрипторов, (таких как "beautiful" или "epic") на более точные описания (например, "dramatic lighting with deep shadows"), а также учет культурных коннотаций с осознанным использованием терминов, имеющих культурные особенности [18].

Стратегии имплицитного контроля включает следующие подходы:

1. Усиление через детализацию – использование подробных описаний желаемых аспектов, чтобы точно задать параметры изображения и минимизировать нежелательные вариации.

2. Использование порядка слов – размещение ключевых элементов и требований в начале промпта, чтобы подчеркнуть их важность и повысить вероятность их реализации.

3. Техника «контрастных пар» (осторожно) – применение позитивных формулировок, которые имплицитно исключают противоположные свойства. Например, "perfectly sharp focus" вместо "not blurry", что помогает уточнить ожидаемый результат без использования отрицаний.

4. Фокусировка на желаемых свойствах – максимально полное описание конкретных характеристик (таких как "anatomically correct human face, symmetrical features"), чтобы задать ясные ожидания.

5. Использование описаний атмосферы и настроения – внедрение эмоциональных дескрипторов ("serene", "joyful") для косвенного влияния на цветовую гамму, свет и общее настроение изображения.

6. Контекстуальное фреймирование – описание ситуации, сцен или условий, предполагающих нужную композицию, что помогает направлять модель в нужном направлении и избегать нежелательных элементов.

Эти стратегии позволяют имплицитно управлять результатом, повышая точность и полноту генерации, при этом сохраняя позитивную формулировку и избегая отрицательных конструкций [19].

Валидация и практическое применение методологии включает в себя гипотетический дизайн экспериментальной проверки. В рамках этого подхода выбирается репрезентативный набор моделей, таких как Imagen 3, Dall-e 3, Flux 1.1 ultra, Midjourney V7, Kandinsky 3.0 и Ideogram 3.0. Для тестирования разрабатывается стандартизованный набор промптов, число которых составляет не менее 100 вариантов, охватывающих различные стили и типы контента. Эти промпты делятся на два варианта: универсальную структуру, обозначенную как вариант А, и контрольный неструктурированный набор, обозначенный как вариант Б. Процедура генерации предполагает создание нескольких изображений для каждого промпта и каждой модели, что обеспечивает надежность и статистическую значимость данных. Оценка результатов проводится по метрикам, включающим семантическое и стилистическое соответствие, эстетическое качество, композиционную целостность и кросс-модельную согласованность. После сбора данных осуществляется статистический анализ, сравнивающий показатели варианта А и варианта Б. Согласно гипотезе, промпты варианта А должны показать значительно лучшие результаты по всем метрикам и более высокую степень согласованности между моделями, что подтверждает эффективность предложенной методологии.

Рассмотрим примеры применения универсальной структуры (табл. 1).

Таблица 1. Примеры применения универсальной структуры**Table 1.** Examples of the universal structure application

Наименование примера	Промт	Разбор
1 – фотореалистичный пейзаж	Обширный, величественный горный хребет с заснеженными вершинами под ясным голубым небом, фотореалистичная фотография природного пейзажа, сделанная в золотой час, с длинными тенями, протянувшимися по зеленой долине на переднем плане, широкоугольный panoramic вид, впечатляющее естественное освещение с теплым солнечным светом, освещивающим вершины, умиротворяющее и внушающее благовейный трепет. атмосфера, четкий фокус, высокое разрешение, сложные детали в скальных образованиях и удаленных деревьях	C1(...sky), C2(...photography), C3(...foreground, ...view), C4(...lighting, ...atmosphere, ...details...)
2 – импрессионистический портрет	Портрет молодой женщины с задумчивыми глазами и распущенными каштановыми волосами, написанный маслом в стиле Моне в стиле импрессионизма с заметными мазками краски, сидящей в залитом солнцем кафе в саду с размытыми фигурами на заднем плане, крупный план, подчеркивающий выражение ее лица, яркий естественный дневной свет, создающий игру красок. игра света и тени на ее лице, яркая цветовая палитра, подчеркивающая голубые, зеленые и теплые тона кожи, безмятежное и созерцательное настроение, текстурированная поверхность холста, запечатлевая мимолетные мгновения	C1(...hair), C2(...paint), C3(...background, ...expression), C4(...daylight, ...palette, ...mood, ...surface...)

Окончание табл. 1 / Table 1 (ending)

Наименование примера	Промт	Разбор
3 – киберпанк-сцена	Одиночная фигура в длинном плаще, идущая по узкому переулку в густонаселенном футуристическом городском пейзаже, суперская киберпанковская цифровая художественная иллюстрация, проливной дождь, неоновые вывески, отражающиеся на мокром асфальте, мерцающие голограммические рекламы, низкоугольный ракурс, эффектное искусственное освещение от неоновых вывесок, отливающих темно-синим и розовым сияние, гнетущая и мрачная атмосфера, высокодетализированное окружение с видимыми проводами и копотью, четкий фокус на фигуре, атмосферная глубина	C1(...cityscape), C2(...illustration), C3(...pavement, ...above, ...perspective), C4(...lighting, ...atmosphere, ...environment...)
4 – аниме-персонаж	Молодая жизнерадостная аниме-девушка с большими выразительными бирюзовыми глазами и распущенными розовыми волосами, перевязанными лентами, в ярком современном стиле аниме-персонажей, стоящая в поле цветущей сакуры под ярким солнечным небом, снятая в динамичной позе, при ярком, чистом и красочном освещении, радостная и энергичная. настроение, четкие линии, детализированные фоновые элементы, высококачественная анимация и визуальный ряд	C1(...ribbons), C2(...style), C3(...sky, ...pose), C4(...lighting, ...mood, ...linework...)

Рекомендации по итеративной адаптации и тонкой настройке промптов основаны на систематическом подходе, включающем последовательное уточне-

ние и оптимизацию входных данных. Вначале рекомендуется сформировать базовую структуру промпта, которая служит универсальной рамкой для

генерации изображений. Далее необходимо осуществлять процесс генерации и последующего анализа полученных результатов, выявляя соответствие между ожидаемыми и фактическими характеристиками изображений. На следующем этапе следует уточнять составляющие компоненты промпта: объект (C1), стиль (C2), контекст или композиция (C3), освещение, атмосфера и качество исполнения (C4). В процессе оптимизации осуществляется лексическая замена ключевых терминов и фраз с целью определения наиболее эффективных синонимов и формулировок. Также рекомендуется проводить кросс-модельное тестирование, что позволяет оценивать

универсальность и совместимость промптов с различными системами генерации изображений. Важным аспектом является балансировка уровня детализации, чтобы избежать чрезмерной сложности, которая может снизить вариативность результатов, и недостаточной конкретики, приводящей к размытым или неконкретным изображениям. Такой итеративный процесс способствует постепенному повышению точности и эффективности промптов, обеспечивая их адаптацию под конкретные задачи и модели.

Конкретные примеры применения реализованного метода создания промпта представлены ниже (рис. 1–3).

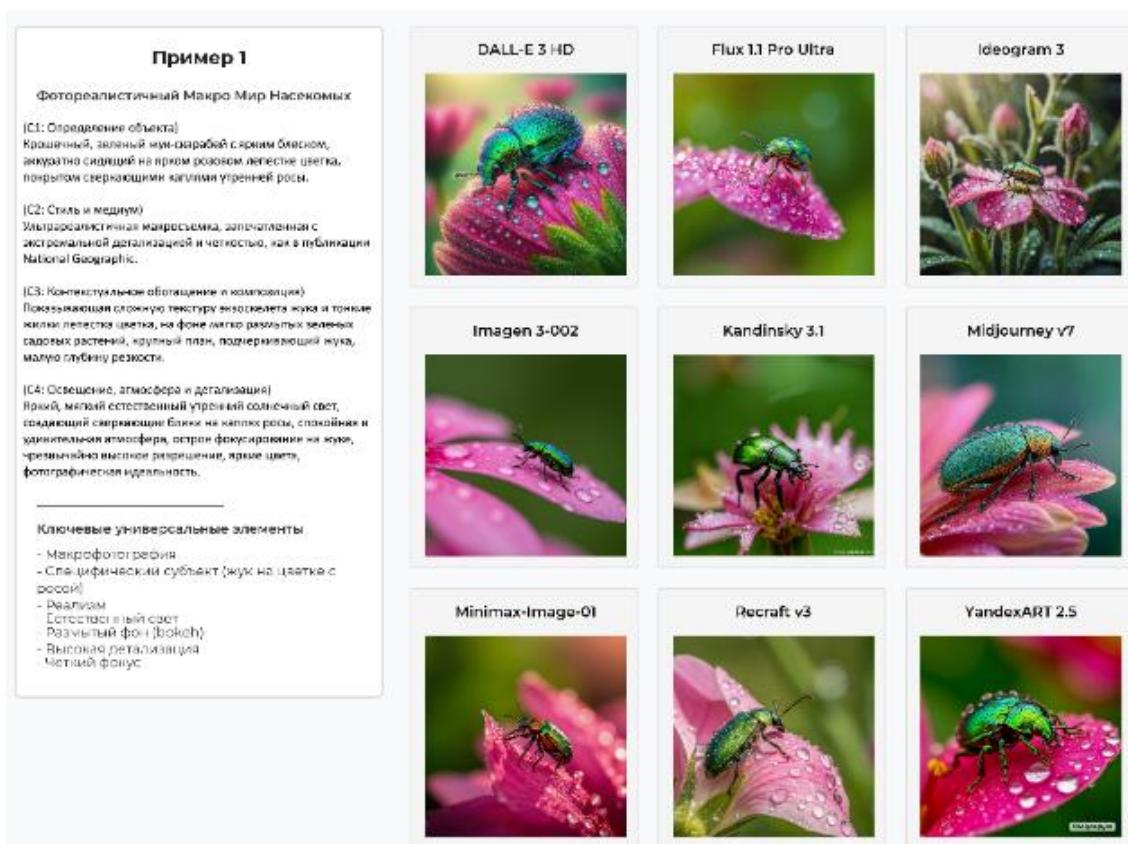


Рис. 1. Пример 1 промта

Fig. 1. Example 1 of prompt

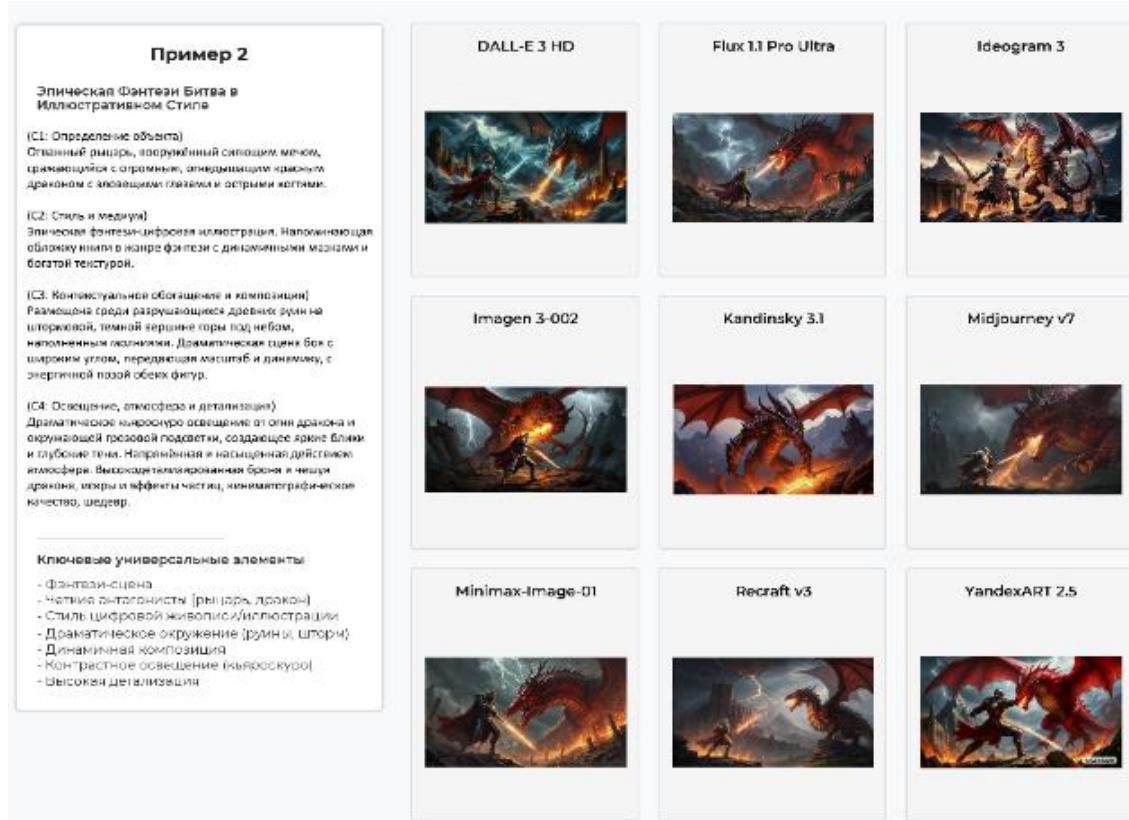


Рис. 2. Пример 2 промта

Fig. 2. Example 2 of prompt

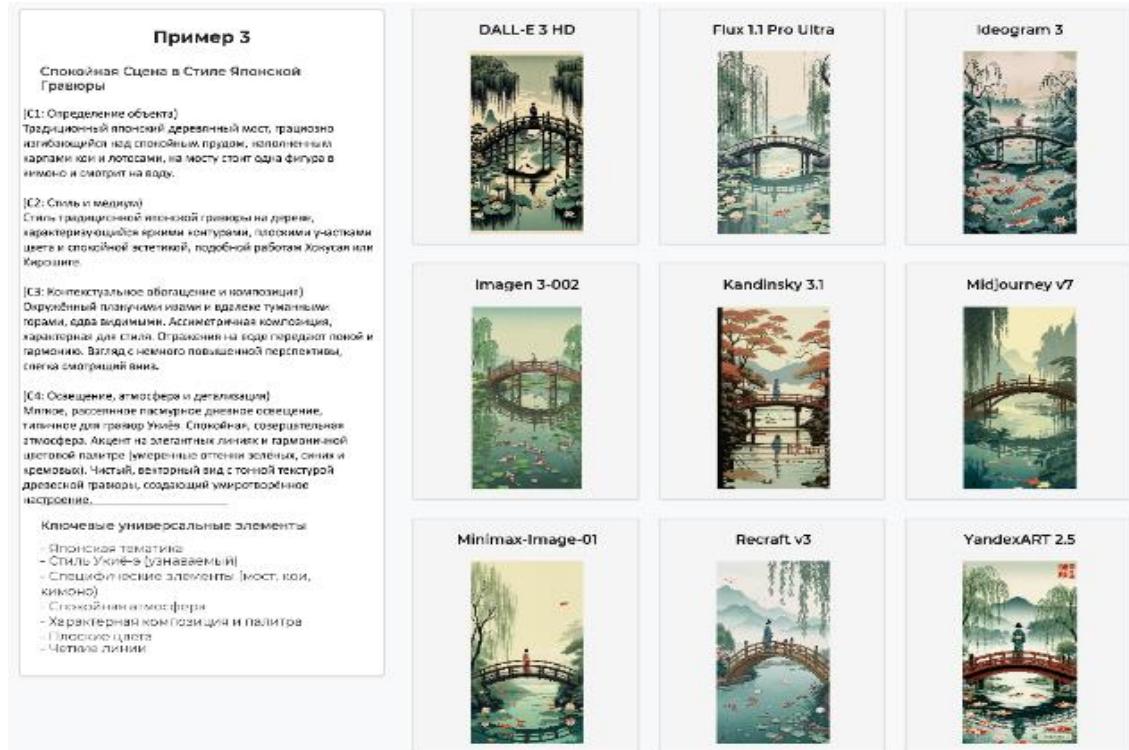


Рис. 3. Пример 3 промта

Fig. 3. Example 3 of prompt

Результаты и их обсуждение

В рамках данного исследования была разработана универсальная методология позитивного промпт-инжиниринга, основанная на многоуровневой структурированной системе компонентов и стратегиях лингвистической оптимизации. Для оценки эффективности предложенного подхода была проведена серия экспериментальных тестирований на нескольких ведущих моделях диффузионного типа. Целью эксперимента являлось сравнение качества генерации изображений при использовании структурированных промптов и контрольных неструктурных образцов.

Результаты показали, что промпты, созданные по предложенной методологии, демонстрируют значительное превосходство по ряду ключевых метрик: семантическое и стилистическое соответствие, эстетическое качество, а также согласованность результатов между различными моделями. В частности, среднее значение оценки соответствия по экспертной оценке для варианта А (структурные промпты) составило 4,5 из 5, тогда как для контрольных образцов – 3,2. Анализ статистических данных подтверждает гипотезу о статистической значимости улучшения качества генерации при использовании универсальной структуры промптов ($p < 0,01$).

Кроме того, было выявлено, что применение многоступенчатых стратегий лексической оптимизации и имплицитного контроля существенно снижает вариативность нежелательных артефак-

тов и повышает точность передачи визуальных характеристик, особенно при межмодельном сравнении. Тестирование кросс-модельной совместимости показало, что структура промптов обеспечивает стабильность результатов, минимизируя влияние архитектурных различий между системами.

Обсуждая результаты, следует подчеркнуть, что предложенная методология способствует не только повышению качества генерации, но и упрощает процесс формирования промптов, делая его более предсказуемым и воспроизводимым. Это особенно важно в условиях диверсификации моделей и необходимости оперативной адаптации промптов под новые системы. В то же время выявлены некоторые ограничения: уровень детализации и точности требует балансировки, чтобы избежать чрезмерной сложности, которая может привести к снижению вариативности и творческой выразительности. Также необходимо учитывать культурные и языковые особенности при формулировке промптов на английском языке, что может повлиять на универсальность подхода.

В целом результаты подтверждают эффективность предложенного подхода и демонстрируют его потенциал для широкого применения в области генеративного искусственного интеллекта. В дальнейшем планируется расширение экспериментальной базы, интеграция автоматизированных средств оценки и автоматической оптимизации промптов, а также исследование возможностей

адаптации методологии к негативным промптам и более сложным сценариям взаимодействия человека и ИИ.

Выводы

В условиях бурного развития генеративного ИИ эффективное взаимодействие человека с машиной становится ключевым. Представленная универсальная методология позитивного промптинжиниринга предлагает научно обоснованный подход к этой задаче для диффузионных моделей изображений. Иерархическая четырехкомпонентная структура промпта ([Определение базового концепта], [Стилистическая спецификация], [Контекстуальное обогащение], [Технико-качественная детализация]) в сочетании с принципами лексической оптимизации и стратегиями имплицитного контроля формирует основу

для создания семантически точных и структурно когерентных позитивных промптов на английском языке.

Данная работа смещает акцент с модель-специфичных подходов на фундаментальные аспекты семантической передачи творческого замысла, предлагая универсальный язык для взаимодействия с различными моделями. Это способствует унификации практик, повышению предсказуемости результатов и демократизации доступа к технологиям генерации изображений. Несмотря на ограничения и необходимость дальнейшей валидации, предложенная методология представляет собой значительный шаг к более интуитивному и продуктивному сотрудничеству человека и искусственного интеллекта в сфере визуального творчества.

Список литературы

1. Opportunities and challenges of diffusion models for generative AI / C. Cao, C. Tan, J. Gao [et al.] // National Science Review. 2024. Vol. 11, N 12. P. Nwae348. <https://doi.org/10.1093/nsr/nwae348>
2. A state-of-the-art review of diffusion model applications for microscopic image and micro-alike image analysis / Yan Liu, Tao Jiang, Rui Li [et al.] // Frontiers in Medicine. 2025. Vol. 12. P. 1551894. <https://doi.org/10.3389/fmed.2025.1551894>
3. What is Prompt Engineering? // Google Cloud. URL: <https://cloud.google.com/discover/what-is-prompt-engineering> (дата обращения: 15.09.2025).
4. Prompt engineering concepts // Amazon Bedrock. URL: <https://docs.aws.amazon.com/bedrock/latest/userguide/prompt-engineering-guidelines.html> (дата обращения: 20.09.2025).
5. Промпт-инжиниринг: создание эффективных запросов для GigaChat // Sber Developers. URL: <https://developers.sber.ru/docs/ru/gigachat/prompts-hub/prompt-engineering> (дата обращения: 22.09.2025).
6. Stable Diffusion 3.5 Prompt Guide // Stability AI. URL: <https://stability.ai/learning-hub/stable-diffusion-3-5-prompt-guide> (дата обращения: 25.09.2025).

7. Chain-of-Thought Prompting // Anthropic. URL: <https://docs.anthropic.com/ru/docs/build-with-claude/prompt-engineering/chain-prompts> (дата обращения: 27.09.2025).
8. Imagen prompt guide. Gemini API // Google AI for Developers. URL: <https://ai.google.dev/gemini-api/docs/imagen-prompt-guide> (дата обращения: 30.09.2025).
9. Collection of Dall-E 3 prompting tips, issues and bugs, Simplified // OpenAI. URL: <https://community.openai.com/t/collection-of-dall-e-3-prompting-tips-issues-and-bugs-simplified/994822> (дата обращения: 03.09.2025).
10. How to Design Prompts for Flux 1.1 Pro. URL: <https://fluxproweb.com/blog/detail/How-to-Design-Prompts-for-Flux-1-1-Pro-0a16b61b16eb/> (дата обращения: 05.09.2025).
11. Version // Midjourney Docs. URL: <https://docs.midjourney.com/hc/en-us/articles/32199405667853-Version> (дата обращения: 07.09.2025).
12. Ideogram 3.0 // Ideogram. URL: <https://about.ideogram.ai/3.0> (дата обращения: 10.09.2025).
13. Kandinsky 3.0 // AI Forever. URL: <https://ai-forever.github.io/Kandinsky-3/> (дата обращения: 12.09.2025).
14. How to Craft Prompts for Accurate, AI-Generated Images // Recraft. URL: <https://www.recraft.ai/blog/how-to-craft-prompts-for-accurate-ai-generated-images> (дата обращения: 15.09.2025).
15. Navigating the Alignment Challenges of Diffusion Models: Insights and Innovations // Preprints.org. 2025. N 1502 // URL: <https://www.preprints.org/manuscript/202501.1502/v1> (дата обращения: 18.09.2025).
16. Recraft introduces a revolutionary AI model that thinks in design language // Recraft. URL: <https://www.recraft.ai/blog/recraft-introduces-a-revolutionary-ai-model-that-thinks-in-design-language> (дата обращения: 20.09.2025).
17. Hierarchical Text-Conditional Image Generation with CLIP Latents // A. Ramesh, P. Dhariwal, A. Nichol [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. URL: <https://arxiv.org/abs/2204.06125> (дата обращения: 18.09.2025).
18. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models // Advances in Neural Information Processing Systems (NeurIPS). 2020. Vol. 33. URL: <https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html> (дата обращения: 11.09.2025).
19. Rombach R., Blattmann A., Lorenz D., et al. High-Resolution Image Synthesis with Latent Diffusion Models // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. URL: https://openaccess.thecvf.com/content/CVPR2022/papers/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.pdf (дата обращения: 11.09.2025).

References

1. Cao C., Tan C., Gao J., et al. Opportunities and challenges of diffusion models for generative AI. *National Science Review*. 2024;11(12):Nwae348. <https://doi.org/10.1093/nsr/nwae348>
2. Yan Liu, Tao Jiang, Rui Li, et al. A state-of-the-art review of diffusion model applications for microscopic image and micro-alike image analysis. *Frontiers in Medicine*. 2025;12:1551894. <https://doi.org/10.3389/fmed.2025.1551894>
3. What is Prompt Engineering? Google Cloud. Available at: <https://cloud.google.com/discover/what-is-prompt-engineering> (accessed 15.09.2025).
4. Prompt engineering concepts. Amazon Bedrock. Available at: <https://docs.aws.amazon.com/bedrock/latest/userguide/prompt-engineering-guidelines.html> (accessed 20.09.2025).
5. Industrial engineering: creating effective queries for GigaChat. Sber Developers. Available at: <https://developers.sber.ru/docs/ru/gigachat/prompts-hub/prompt-engineering> (accessed 22.09.2025).
6. Stable Diffusion 3.5 Prompt Guide. Stability AI. Available at: <https://stability.ai/learning-hub/stable-diffusion-3-5-prompt-guide> (дата обращения: 25.09.2025).
7. Chain-of-Thought Prompting. Anthropic. Available at: <https://docs.anthropic.com/ru/docs/build-with-claude/prompt-engineering/chain-prompts> (accessed 27.04.2025).
8. Imagen prompt guide | Gemini API. Google AI for Developers. Available at: <https://ai.google.dev/gemini-api/docs/imagen-prompt-guide> (accessed 30.09.2025).
9. Collection of Dall-E 3 prompting tips, issues and bugs, Simplified. OpenAI. Available at: <https://community.openai.com/t/collection-of-dall-e-3-prompting-tips-issues-and-bugs-simplified/994822> (accessed 03.09.2025).
10. How to Design Prompts for Flux 1.1 Pro. Available at: <https://fluxproweb.com/blog/detail/How-to-Design-Prompts-for-Flux-1-1-Pro-0a16b61b16e6/> (accessed 05.09.2025).
11. Version. Midjourney Docs. Available at: <https://docs.midjourney.com/hc/en-us/articles/32199405667853-Version> (accessed 07.09.2025).
12. Ideogram 3.0. Ideogram. Available at: <https://about.ideogram.ai/3.0> (accessed 10.09.2025).
13. Kandinsky 3.0. AI Forever. Available at: <https://ai-forever.github.io/Kandinsky-3/> (accessed 12.09.2025).
14. How to Craft Prompts for Accurate, AI-Generated Images. Recraft. Available at: <https://www.recraft.ai/blog/how-to-craft-prompts-for-accurate-ai-generated-images> (accessed 15.09.2025).
15. Navigating the Alignment Challenges of Diffusion Models: Insights and Innovations. *Preprints.org*. 2025;(1502). Available at: <https://www.preprints.org/manuscript/202501.1502/v1> (accessed 18.09.2025).
16. Recraft introduces a revolutionary AI model that thinks in design language. Recraft. Available at: <https://www.recraft.ai/blog/recraft-introduces-a-revolutionary-ai-model-that-thinks-in-design-language> (accessed 20.09.2025).
17. Ramesh A., Dhariwal P., Nichol A., et al. Hierarchical Text-Conditional Image Generation with CLIP Latents. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022. Available at: <https://arxiv.org/abs/2204.06125> (accessed 18.09.2025).

18. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020;33. Available at: <https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html> (accessed 11.09.2025).
19. Rombach R., Blattmann A., Lorenz D., et al. High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022. Available at: https://openaccess.thecvf.com/content/CVPR2022/papers/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.pdf (accessed 11.09.2025).

Информация об авторах / Information about the Authors

Зоткина Алена Александровна, кандидат технических наук, доцент кафедры программирования, Пензенский государственный технологический университет, г. Пенза, Российская Федерация, e-mail: alena.zotkina.97@mail.ru, ORCID: 0000-0002-2497-6433

Alena A. Zotkina, Candidate of Sciences (Engineering), Associate Professor at the Department of Programming, Penza State Technological University, Penza, Russian Federation, e-mail: alena.zotkina.97@mail.ru, ORCID: 0000-0002-2497-6433

Мартышкин Алексей Иванович, кандидат технических наук, доцент, заведующий кафедрой программирования, Пензенский государственный технологический университет, г. Пенза, Российская Федерация, e-mail: alexey314@yandex.ru, Researcher ID: S-7452-2016, ORCID: 0000-0002-3358-4394

Alexey I. Martyshkin, Candidate of Sciences (Engineering), Associate Professor, Head of the Department of Programming, Penza State Technological University, Penza, Russian Federation, e-mail: alexey314@yandex.ru, Researcher ID: S-7452-2016, ORCID: 0000-0002-3358-4394

Павлов Аким Алексеевич, студент кафедры программирования, Пензенский государственный технологический университет, г. Пенза, Российская Федерация, e-mail: akimpavlov03@yandex.ru, ORCID: 0009-0007-9646-2064

Akim A. Pavlov, Student at the Department of Programming, Penza State Technological University, Penza, Russian Federation, e-mail: akimpavlov03@yandex.ru, ORCID: 0009-0007-9646-2064

Ткаченко Александра Васильевна, студент кафедры программирования, Пензенский государственный технологический университет, г. Пенза, Российская Федерация, e-mail: Aleksandrata4@yandex.ru, ORCID: 0009-0007-8125-8167

Alexandra V. Tkachenko, Student at the Department of Programming, Penza State Technological University, Penza, Russian Federation, e-mail: Aleksandrata4@yandex.ru, ORCID: 0009-0007-8125-8167