

*Курочкин Вадим Сергеевич, 2 курс магистратуры, Цифровые технологии  
обработки данных, РТУ МИРЭА*

*Андреева Арина Александровна, 2 курс магистратуры, Цифровые  
технологии обработки данных, РТУ МИРЭА*

*Гетман Владислав Антонович, 2 курс магистратуры, Цифровые  
технологии обработки данных, РТУ МИРЭА*

## **ИСПОЛЬЗОВАНИЕ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ СОЗДАНИЯ ДИЗАЙН-КОНЦЕПТА ТИРАЖИРУЕМОГО ИЗОБРАЖЕНИЯ НА ПРОИЗВОДСТВЕ**

*В данной статье рассматривается разработка интеллектуальной информационной системы для тиражируемого производства, основным функционалом которой является генерация дизайна изображения по текстовому запросу. Предлагаемая система направлена на автоматизацию процесса создания визуального контента, значительно ускоряя и упрощая работу дизайнеров и художников. Основной целью системы является генерация визуально привлекательных и соответствующих текстовым запросам изображений с помощью передовых методов машинного обучения. Предлагаемое решение включает в себя разработку программных модулей: клиентское приложение для взаимодействия с системой, модуль работы с генератором изображений.*

**Ключевые слова:** *Генерация изображений, Нейронные сети, Машинное обучение, Тиражируемое производство, Автоматизация дизайна, Визуальный контент, Текстовые запросы, DALL-E, GPT, Глубокое обучение*  
*This article discusses the development of an intelligent information system for replicated production, the main functionality of which is the generation of an image design based on a text query. The proposed system is aimed at automating the process of creating visual content, significantly speeding up and simplifying the work of designers and artists. The main purpose of the system is to generate visually*

*appealing and text-relevant images using advanced machine learning techniques. The proposed solution includes the development of software modules: a client application for interacting with the system, a module for working with an image generator.*

**KEYWORDS:** *Image Generation, Neural Networks, Machine Learning, Replicated Production, Design Automation, Visual Content, Text Queries, DALL-E, GPT, Deep Learning*

## **Введение**

Привлекательный дизайн упаковки конфет имеет значительное влияние на решение потребителей при покупке товара. Это означает, что эффектные и оригинальные дизайны обёрток могут помочь компаниям привлечь больше клиентов и увеличить свою прибыль. В настоящее время одним из эффективных подходов к разработке новых дизайнов является использование компьютерных технологий для создания виртуальных прототипов. В частности, нейронные сети и искусственный интеллект хорошо зарекомендовали себя как способ качественной обработки информации и генерации новой на основе входных данных. Говоря про актуальность исследования, генерация визуального контента является важной и востребованной задачей во многих областях, включая рекламу, электронную коммерцию, издательскую деятельность и дизайн. Традиционно создание изображений и иллюстраций требует значительных временных и трудовых затрат со стороны профессиональных художников и дизайнеров. Однако растущие потребности в визуальном контенте, особенно в условиях цифровой экономики, диктуют необходимость более эффективных и автоматизированных решений.

## **Цель исследования**

Целью работы является создание прототипа ИТ системы для сокращения сроков разработки дизайна изображения на тиражируемом производстве путём использования нейронных сетей.

Для достижения поставленной цели была сформулирована исследовательская гипотеза, что современные алгоритмы глубокого обучения позволяют в значительной степени ускорить и упростить процесс создания дизайна.

Для проверки исследовательской гипотезы были поставлены исследовательские вопросы (ИВ):

ИВ1: изучить этапы технологического производства, выделить ограничения, накладываемые оборудованием;

ИВ2: изучить популярные архитектуры генераторов изображений на основе текстового запроса;

ИВ3: разработка программного обеспечения.

## **1. Предметная область и основные определения**

### **Этапы производственного процесса**

Для ответа на ИВ1 необходимо рассмотреть этапы производства. В данном случае за пример было взято изготовление оберток для конфетных изделий.

В общем случае процесс изготовления обёртки можно разделить на 4 этапа:

- изготовление материала;
- нанесение декоративных элементов;
- резка и формирование размера и формы;
- упаковка.

В зависимости от материала технология производства может сильно отличаться. Например, упаковку из вторсырья делают из бумажной массы, которую сначала очищают и сортируют. Далее её проклеивают и окрашивают в однородный цвет. В конце концов массу отливают в формы, сушат и снимают для дальнейшей сушки (Рисунок 1) [1].

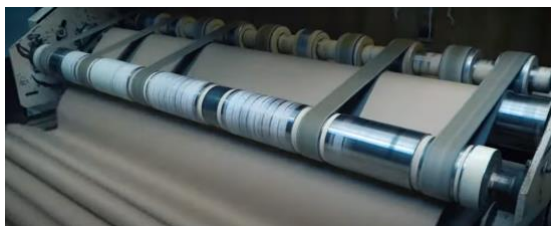


Рисунок 1 — Производство бумаги из вторсырья

Если речь идёт о полимерной плёнке, то методов изготовления тоже несколько. Один из таких — экструзионный. В простейшем виде такой метод представляет из себя выдавливание расплавленной полимерной массы через специальный формирующий инструмент. Когда производство сталкивается с многослойной плёнкой, то используется несколько экструдеров, загруженные разными типами сырья. Экструдеры работают в одну сложную по конструкции головку, оформляющую несколько слоёв с разными свойствами (Рисунок 2) [2].



Рисунок 2 — Плоскощелевой экструдер для производства пятислойной стрейч-пленки

Нанесение декора также отличается вариативностью не только в зависимости от материала, но и в рамках одного. Из основных способов можно выделить УФ лакирование прозрачными и глиттерными лаками, тиснение фольгой, вырубку (Рисунок 3).



Рисунок 3 — Флексографическая машина УТ-2-600

После нанесения декора материал режется на отдельные листы или полосы, которые затем формируются в обёртки нужного размера и формы. Этот этап может выполняться с использованием автоматизированных машин

или ручным способом, в зависимости от масштаба производства. Готовый рулон упаковочной плёнки изображён на Рисунке 4.



Рисунок 4 — Рулон упаковочной плёнки

Последний этап — это упаковка и отправка на кондитерские фабрики.

### **Ограничения, накладываемые оборудованием**

Исходя из производственного процесса, можно выделить определенные ограничения для результата работы системы.

Упаковочное оборудование имеет определенные диапазоны размеров, которые оно может обрабатывать эффективно. Для изображений это означает, что они должны быть достаточно большими, чтобы быть видимыми на обёртке, но не слишком большими, чтобы не превышать размеры упаковочного материала или не помещаться в область печати.

Изображения также должны соответствовать пропорциям упаковки, чтобы не искажаться при печати или при обёртке. Например, леденцы имеют эллипсоидную форму, а значит рисунок может быть искажён после упаковки.

Стоит учитывать и наличие безопасных зон, где не должны находиться важные элементы дизайна, чтобы они не были обрезаны во время процесса обработки, обрезки или упаковки [3].

Цветовая палитра — важная часть разработки дизайна. Нейронная сеть может генерировать огромное количество различных образов и рисунков в большом разнообразии цветов. Но далеко не каждое изображение может быть воспроизведено на производстве обёрток.

В зависимости от метода печати на цветовую палитру накладываются определенные ограничения. Некоторые методы не могут воспроизвести определенные цвета и оттенки. Например, СМУК печать не может воспроизводить яркие насыщенные цвета, а также некоторые специфические цвета, как золотистый [4]. Всё разнообразие палитры может предоставить только цифровая печать, но она является высоко затратной при больших тиражах, а также не подходит для больших форматов.

Некоторые материалы также могут ограничивать доступные цвета. Например, на коричневой крафт-бумаге рисунки будут «приглушённые» или вообще иметь палитру только из оттенков коричневого.

Требования к разрешению и детализации обычно определяются возможностями производства и техническими характеристиками используемого оборудования.

Для качественного изображения на упаковке разрешение не должно быть менее 300 DPI, иначе рисунок сильно потеряет в качестве из-за пикселизации [3] [5].

В обратном же случае, когда разрешение изображения слишком большое и содержит мелкие детали, печатное оборудование может не справиться с такой детализацией, и некоторые части рисунка не распечатаются. В некоторых случаях даже мелкие шрифты могут быть неразличимы для чтения.

Данные ограничения возможно наложить на разных этапах работы системы. Один из вариантов – использование ключевых слов в промте. Но такой способ потребует более тщательного препроцессинга датасета для обучения и сложной разметки изображений. Гораздо проще наложить данные ограничения после генерации, используя простые алгоритмы цветокоррекции, обрезки под нужные размеры и сжатия до определенной детализации.

## **Использование нейронных сетей для генерации изображений**

Чтобы ответить на ИВ2 необходимо изучить архитектуры некоторых реализаций нейронных сетей, способных генерировать изображения по тексту. Такими сетями являются DALL-E и Stable Diffusion.

DALL-E состоит из двух частей — GPT и VQGAN. Первая имеет архитектуру декодера от трансформера, которая способна видеть связи между токенами текста и токенами изображения. Вторая часть сети способна декодировать последовательность токенов, полученную от GPT и преобразовывать в картинку, соответствующую входному тексту. На Рисунке 5 изображена упрощенная схема работы DALL-E 2 [6].

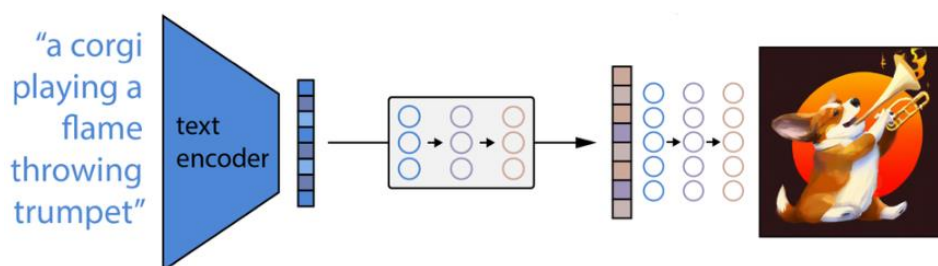


Рисунок 5 — Упрощенная архитектура сети DALL-E 2

Stable Diffusion имеет определенные отличия в своей работе. Первый блок сети – это компонент, который преобразует текстовую информацию в цифровой вид, который передаёт заложенный в текст смысл. Здесь также используется специальная языковая модель трансформер. Далее информация передается к генератору изображения, который состоит из двух частей.

Первый компонент генератора — Image Information Creator, а второй декодер автоэнкодера. Благодаря такому генератору Stable Diffusion значительно подняла качество изображений по сравнению с предыдущими моделями.

В Image Information Creator используется процесс, называемый диффузией, который воспроизводится с помощью сети UNet и планировщика. Имея эмбединги токенов, описывающие введенный текст, и случайный шум, процесс создаёт массив информации, который декодер изображения использует для рисования готового изображения.

Диффузия выполняется в несколько шагов, каждый из которых работает с входным массивом и создаёт ещё один массив, который лучше представляет введённый текст, а вся визуальная информация модели собирается из всех изображений, на которых была обучена модель [7].

Схема Stable Diffusion представлена на Рисунке 6.

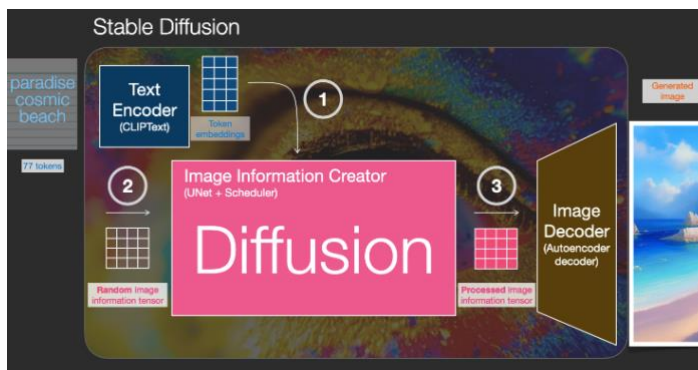


Рисунок 6 — Схема архитектуры сети Stable Diffusion

Приведенные примеры демонстрируют лишь одни из первых успешных попыток генерации изображений по тексту. На данный момент архитектуры моделей нейронных сетей ушли далеко вперед как по быстродействию, так и по результатам.

Для разрабатываемой системы предлагается использовать сеть DALL-E. Она имеет несложную архитектуру, но при этом демонстрирует впечатляющие результаты.

## 2. Предлагаемое решение по сокращению сроков разработки дизайна на тиражируемом производстве

Исходя из информации, полученной при анализе предметной области была разработана модель интеллектуальной информационной системы в нотации IDEF0 и DFD.

Система в нотации IDEF0 изображена на Рисунке 7.

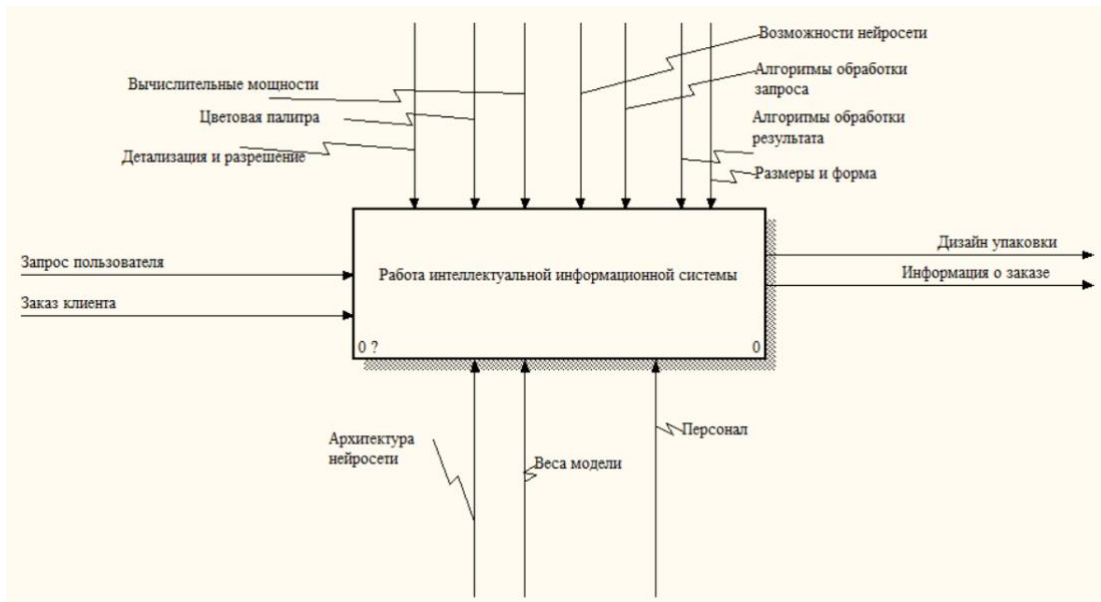


Рисунок 7 — Модель в нотации IDEF0

Декомпозиция состоит из трёх блоков: обработка запроса пользователя, работа нейросети, работа с результатом генерации нейросети.

На первом этапе подаётся запрос пользователя и данные заказа клиента. После обработки запрос в виде закодированного промта поступает на вход нейросети. Нейросеть генерирует изображение и подаёт его на вход блоку «Работа с результатом генерации нейросети», где результат сохраняется и, при необходимости, создаётся новый запрос, который снова поступает на вход первому блоку. Декомпозиция представлена на Рисунке 8.

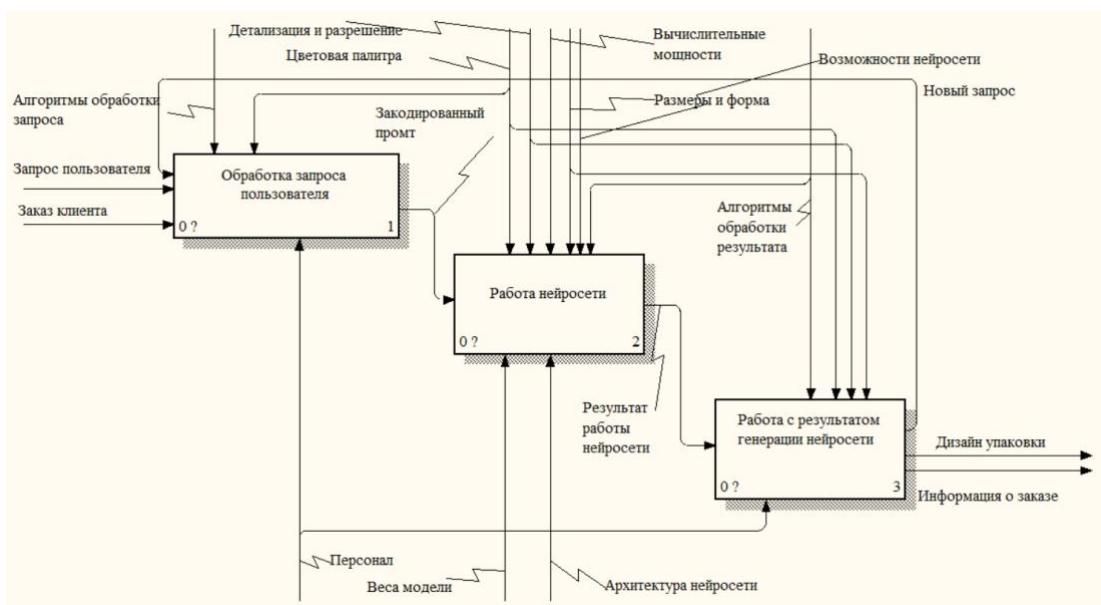


Рисунок 8 — Декомпозиция модели IDEF0

Декомпозиция блока «Обработка запроса пользователя» (Рисунок 9) состоит из введения запроса и форматирования, сохранения в базу данных, кодирование промта.

Форматирование представляет из себя базовые манипуляции с текстом для удобного сохранения и дальнейшей работы с ним. После этого запрос и новые данные заказа сохраняются в базу данных. На последнем этапе текст кодируется для того, чтобы его могла принять на себя нейронная сеть.

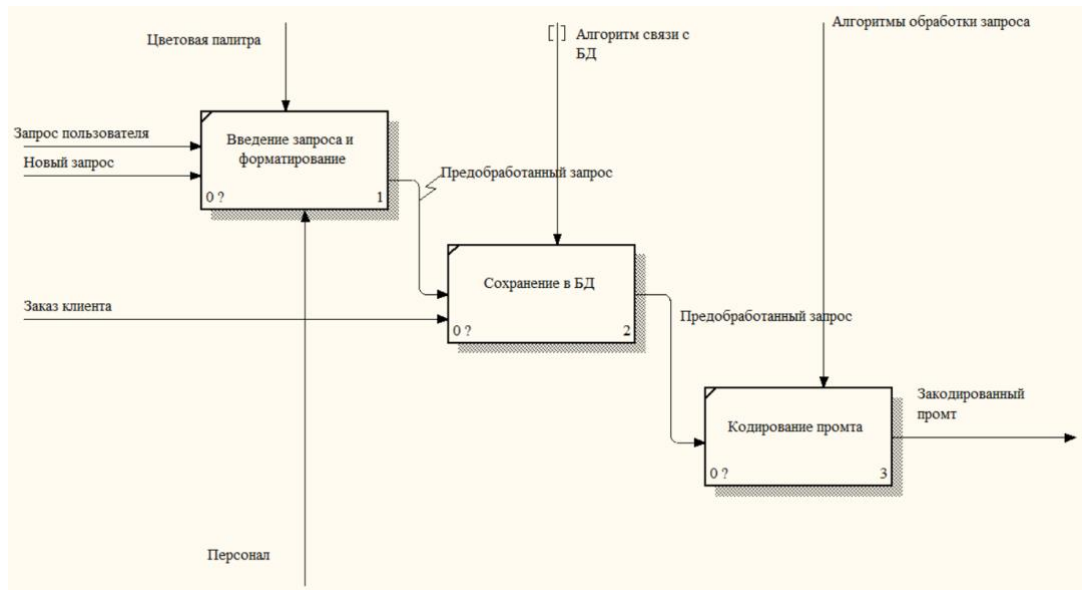


Рисунок 9 — Декомпозиция «Обработка запроса пользователя» IDEF0

Работа нейросети состоит из 3 базовых этапов (Рисунок 10): сопоставление токенов текста со скрытым представлением изображения, работа декодера сети, постобработка изображения.

Закодированный текст GPT использует для генерации последовательности чисел для изображения. Далее эти числа преобразуются в тензор изображения с помощью декодера. На последнем этапе проводится постобработка рисунка.

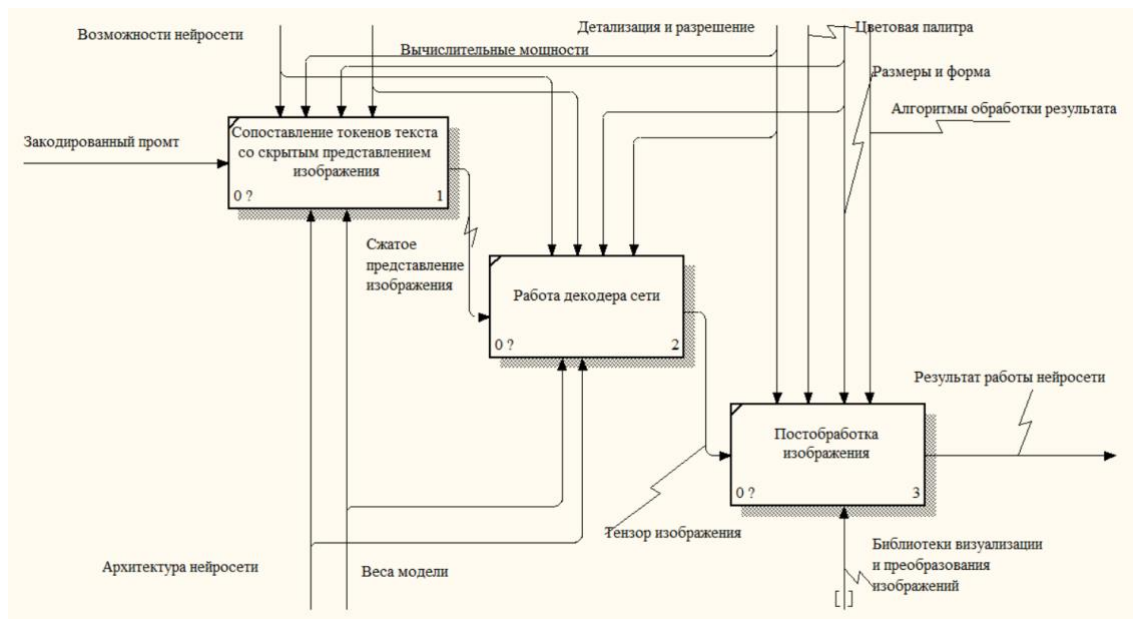


Рисунок 10 — Декомпозиция «Работа нейросети» IDEF0

Декомпозиция блока «Работа с результатом генерации нейросети» (Рисунок 11) включает: внесение корректировок, демонстрация результата, формирование заказа на производство и сохранение и базу данных.

После внесения корректировок в результат может потребоваться создание нового запроса и регенерация, иногда данный этап не включает серьёзных правок. Во время демонстрации результата заказчик также может потребовать внести правки. Если результат соответствует всем необходимым требованиям, то он сохраняется и формируется полноценный заказ на производство.

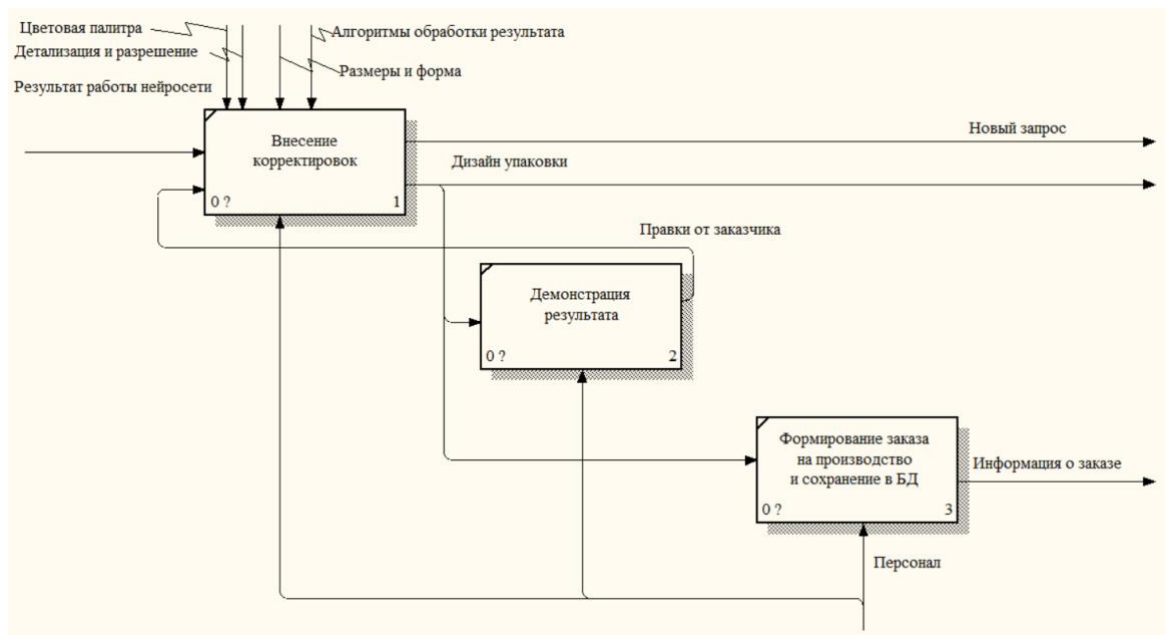


Рисунок 11 — Декомпозиция «Работа с результатом генерации нейросети»  
IDEF0

Модель в нотации DFD изображена на Рисунке 12.

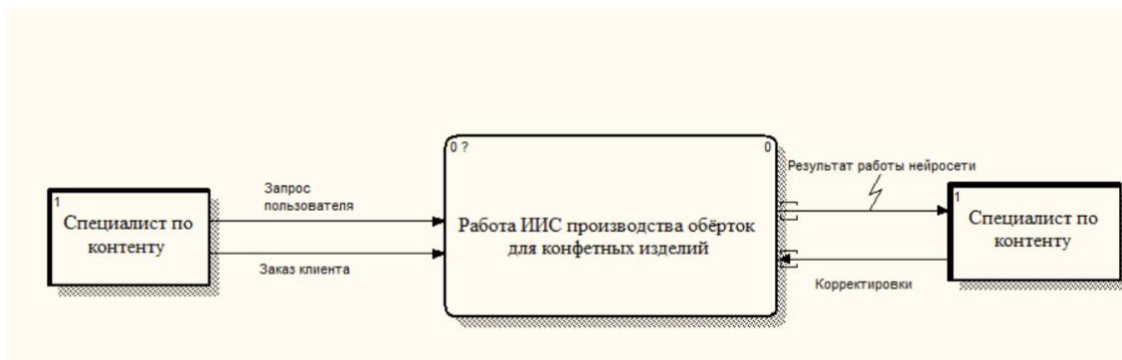


Рисунок 12 — Модель в нотации DFD

Декомпозиция модели в нотации DFD изображена на Рисунке 13. Вся информация о заказе, промт и требования сохраняются в базе данных. Также оттуда загружаются данные для обработки и формирования заказа. Все корректировки сохраняются в БД. Готовый дизайн отправляется во внутреннее хранилище, а в базе хранится числовой код изображения.

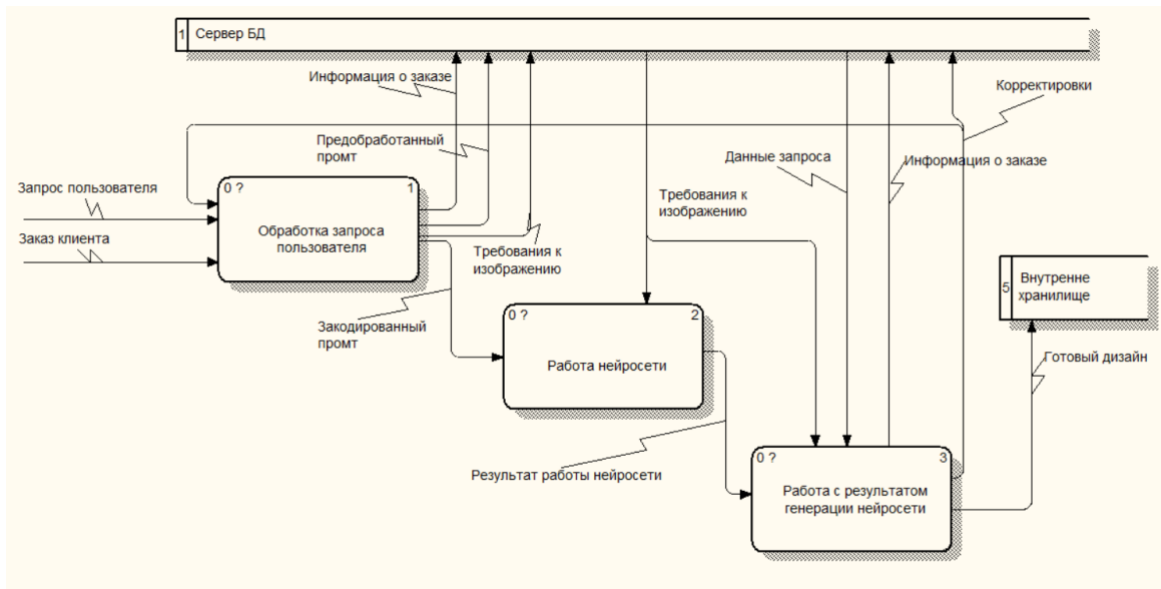


Рисунок 13 — Декомпозиция в нотации DFD

Декомпозиция процесса «Обработка запроса пользователя» в нотации DFD изображена на Рисунке 14.

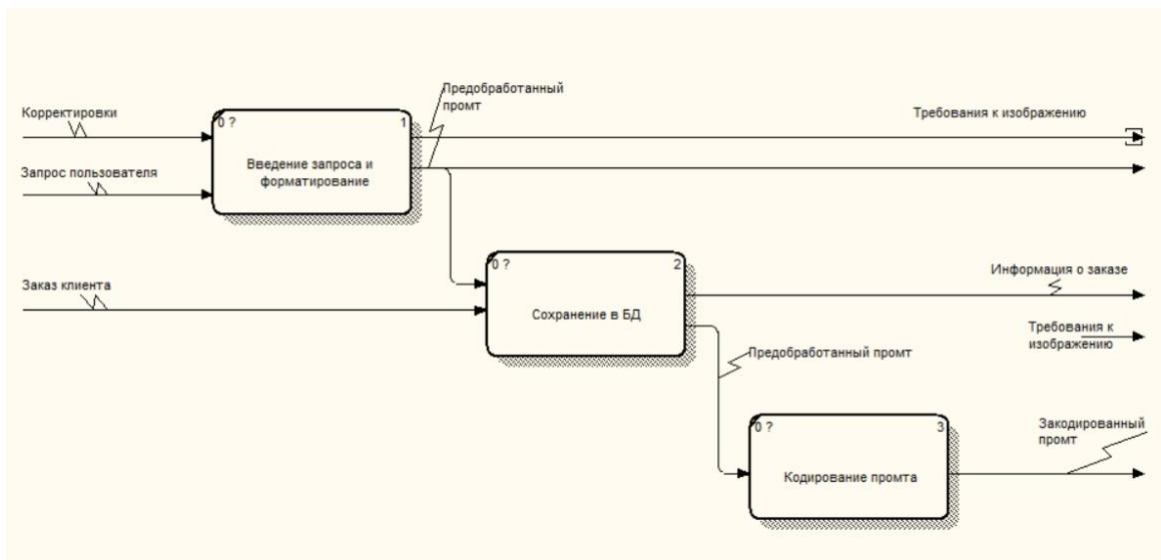


Рисунок 14 — Декомпозиция процесса «Обработка запроса пользователя» в нотации DFD

Декомпозиция процесса «Работа нейросети» в нотации DFD изображена на Рисунке 15.

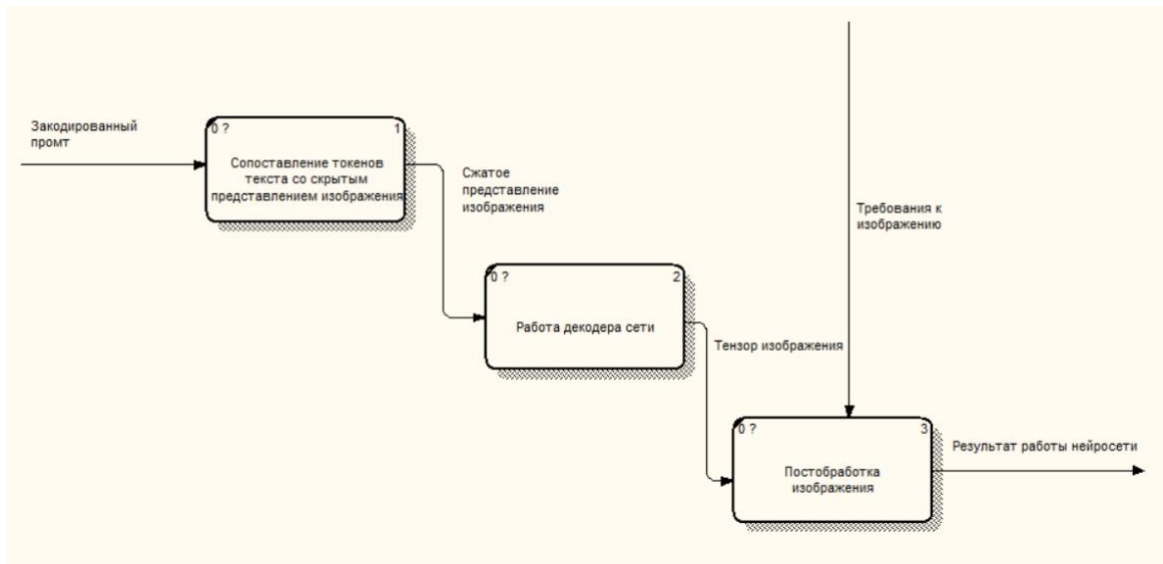


Рисунок 15 — Декомпозиция процесса «Работа нейросети» в нотации DFD

Декомпозиция процесса «Работа с результатом генерации нейросети» в нотации DFD изображена на Рисунке 16.

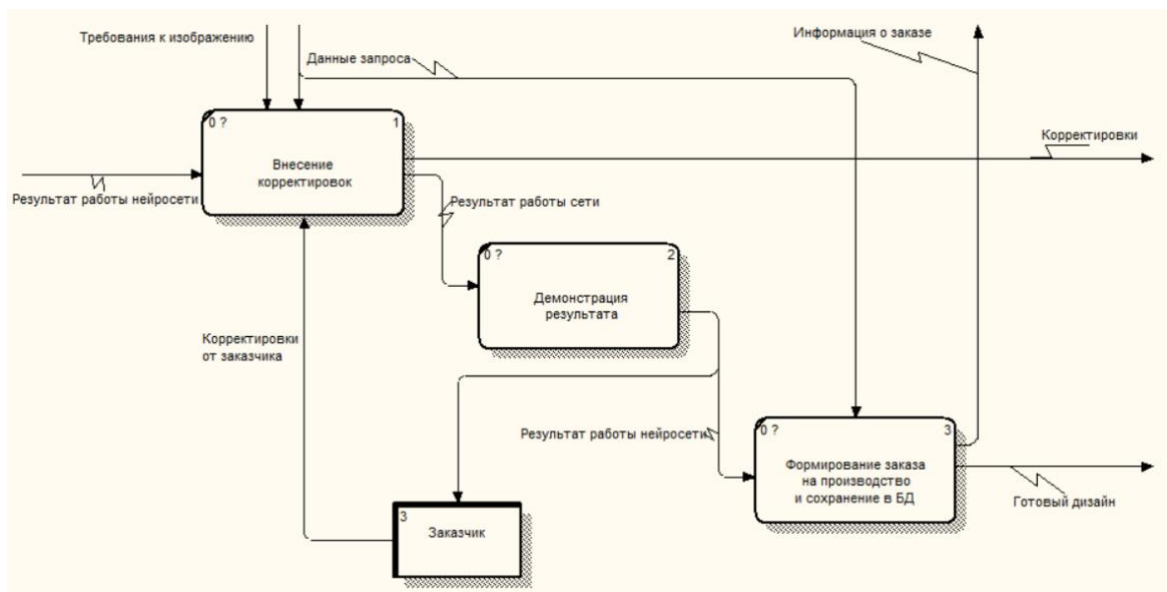


Рисунок 16 — Декомпозиция процесса «Работа нейросети» в нотации DFD

Таким образом, система позволяет намного быстрее формировать дизайн обёртки, даже с участием одного специалиста на заказ.

### 3. Разработка прототипа интеллектуальной информационной системы

В результате анализа предметной области и используя модели в нотациях IDEF0 и DFD были составлены функциональные требования к системе. Требования к функциям и задачам представлены в Таблице 1.

Таблица 1 — Требования к функциям, выполняемым системой

<b>Функция</b>	<b>Задача</b>
Функция обработки запросов	Чтение запроса пользователя
	Обработка запроса для дальнейшей работы
Функция вывода результата	Вывод результата работы
	Возможность сохранить результат
	Перегенерация результата
Функция сохранения результатов работы нейронной сети	Сохранение изображения в базу данных
Функция информирования о сбоях	Предоставление подробной информации о сбое и возможном решении проблемы
Функция логирования данных	Своевременное составление и сохранение логов всех процессов
Функция преобразования текста в изображение	Принятие на вход обработанного запроса и генерация по нему изображения

### **Сбор данных и обучение нейросети**

Для обучения нейросети, которая будет использоваться в интеллектуальной информационной системе, требуется большое количество изображений обёрток и упаковочных рисунков с разметкой.

Источником набора данных была выбрана платформа Adobe Stock. На ней публикуются работы дизайнеров и фотографов, которые к тому же ещё имеют текстовое сопровождение, описывающее изображение.

Для сбора данных был написан парсер, который сохраняет картинку в формате “jpg” и его описание в “txt” файл. После запуска скрипта было скачано 10000 изображений с рисунками на упаковках с различным декором, а также был наполнен файл с их текстовыми описаниями.

После ручной очистки набора данных осталось около 8500 изображений. Для обучения нейронной сети этого также было недостаточно. Были произведены некоторые аугментации над файлами с целью расширения набора данных, в результате чего количество изображений было увеличено до 25500.

Создание и обучение модели производилось с помощью фреймворка машинного обучения PyTorch, который используется для решения различных задач, включая тензорные вычисления и обработка естественного языка.

Для части сети, функция которой декодировать латентное представление входящего текста, была использована предобученная модель с архитектурой VQGAN, которая уже способна качественно «разжимать» изображение. Веса GPT части модели следовало обучить вручную.

Обучение модели в сумме заняло около 24 часов. Среднее значение функции потерь на конец обучения равнялось 2.8.

### **Тестирование работы модуля нейронной сети**

При тестировании работы нейронной сети использовался графический процессор T4 GPU предоставленный платформой Google Colab. При таких вычислительных мощностях на генерацию 9 изображений потребовалось около 20 секунд.

Разработка самой интеллектуальной информационной системы и её тестирование производилось на ноутбуке. Обработку запросов и генерацию изображений выполнял центральный процессор AMD Ryzen 5 3500U с

частотой 2.10 GHz. На загрузку архитектуры и весов моделей ушло 14 секунд. Генерация одного изображения занимает примерно 1 минуту и 50 секунд, что значительно уступает использованию графического процессора от Google Colab.

На Рисунке 17 предоставлен результат работы генератора по запросу “Japanese style gold leaf plant on black background, modern, art deco background pattern. Design for textures, posters, wrapping paper, gift” (с англ. «Растение в японском стиле с золотыми листьями на черном фоне, современный фон в стиле арт-деко. Дизайн текстур, плакатов, оберточной бумаги, подарков»). На Рисунке 18 с запросом “Magical black and gold glitter background. Gift boxes design.” (с англ. «Волшебный черно-золотой фон с блестками. Дизайн подарочных коробок.»). Рисунок 19 демонстрирует результат работы с промптом “Pink and white packaging squares hearts flowers” (с англ. «Розовая и белая упаковка квадратики сердечки цветы»).

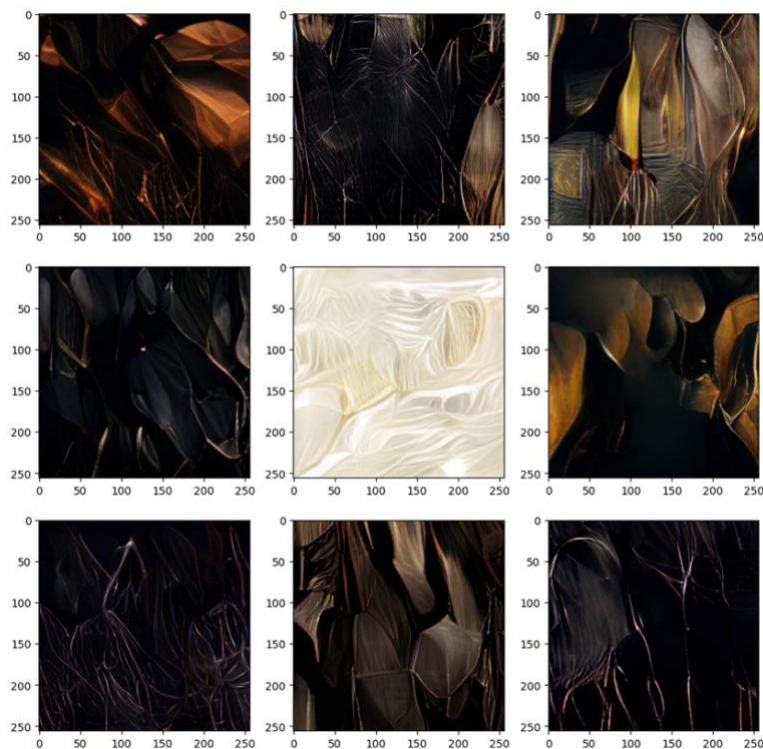


Рисунок 17 — Результат работы генератора по первому запросу

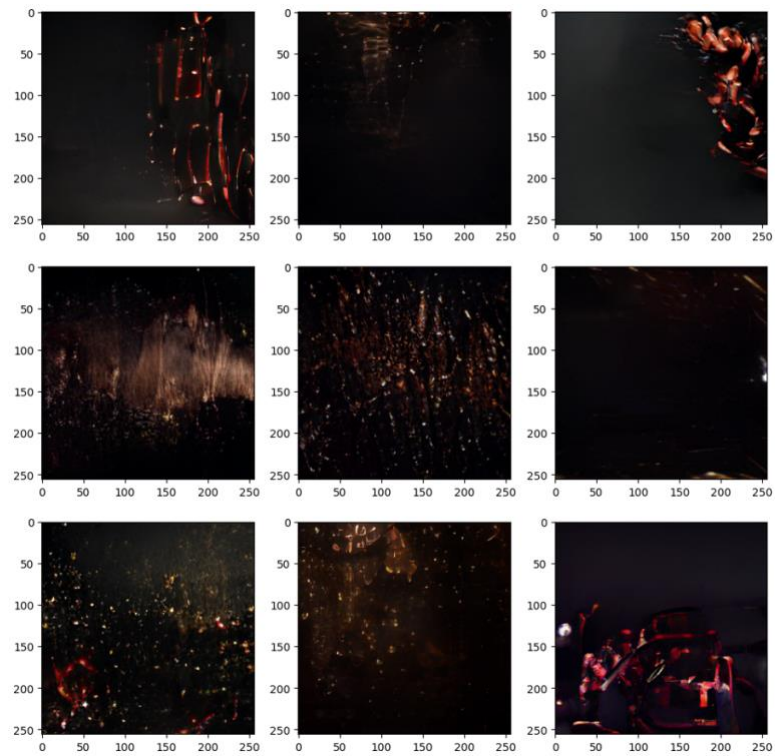


Рисунок 18 — Результат работы генератора по второму запросу

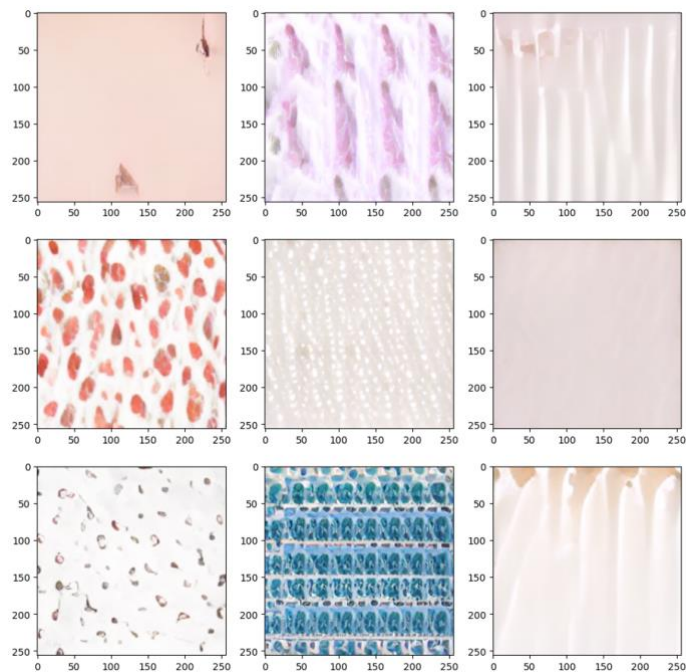


Рисунок 19 — Результат работы генератора по третьему запросу

По результату работы с разными запросами можно сделать вывод, что нейронная сеть хорошо распознает цвета, указанные в промте, а также имеет общее представление о «бесшовности» рисунка.

Алгоритм по-прежнему не может генерировать конкретные образы чего-либо, а предоставляет только возможную текстуру или фон для будущего дизайна. Для улучшения качества генератора требуется больше времени обучения, большее количество размеченных изображений, а также использование графических процессоров с большими вычислительными мощностями.

## Демонстрационное приложение

Для демонстрации работы модулей системы было написано простое приложение с клиентским интерфейсом (Приложение А, Б). Оно включает в себя только некоторые основные функции системы:

- авторизация пользователя (Рисунок 20);
- вывод списка доступных заказов (Рисунок 21);
- просмотр информации о заказе;
- ввод промта и требований к изображению (Рисунок 22);
- генерация изображения посредством нейронной сети;
- просмотр результата работы нейросети (Рисунок 23);
- сохранение результата работы нейросети;
- просмотр предыдущих результатов;
- регенерация результата.

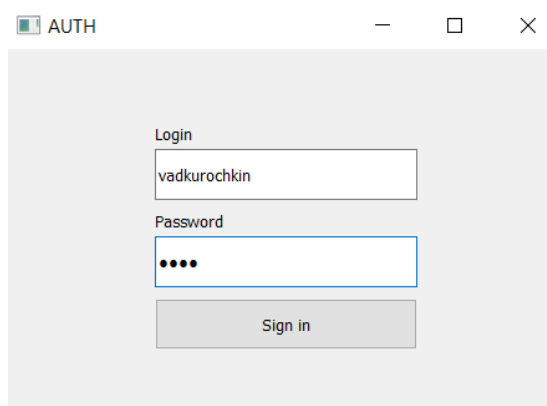


Рисунок 20 — Окно авторизации пользователя

		Client	Status	Date	Description	Full Text
1	Accept order	Zhuravlev	created	2024-05-04	flowers white red stripes	Show text
2	Accept order	Zhuravlev	finished	2024-05-04	hand-painted breaking bread	Show text
3	Accept order	Bushueva	created	2024-05-09	flowers, cats, abstract figures	Show text

Рисунок 21 — Список всех клиентских заказов

Logged in

Prompt: Abstract pattern collection with pink color. Decorative abstract horizontal banher v

Requirements:

- Length: 256
- Width: 256
- Color Palette: RGB
- Resolution: 150

GENERATE

Рисунок 22 — Ввод запроса и требований

Logged in

Prompt: Abstract pattern collection with pink color. Decorative abstract horizontal banher v

Requirements:

- Length: 256
- Width: 256
- Color Palette: RGB
- Resolution: 150

GENERATE

Regenerate

Save and Finish

Previous

1	49145
2	73540
3	56976

Рисунок 23 — Демонстрация результата и инструменты работы с ним

Разработанное демонстрационное приложение выполняет только основные функции. Помимо вышеперечисленных возможностей конечный вариант системы может включать такие функции, как регистрация новых заказов, добавление данных о специалистах, встроенный графический редактор и т.д.

## **Полученные результаты.**

В результате проведенного исследования были найдены ответы на все поставленные исследовательские вопросы.

Была проанализирована предметная область, а именно технологический процесс производства для выявления ограничений, накладываемых оборудованием. Это послужило ответу на ИВ1. Для ответа на ИВ2 были рассмотрены две реализации генераторов изображений по текстовому запросу для выбора подходящей под систему.

После анализа предметной области и проектирования системы, были разработаны программные модули для ответа на ИВ3.

## **Заключение.**

В данной статье была рассмотрена интеллектуальная информационная система, предназначенная для генерации дизайна изображений по текстовому запросу в условиях тиражируемого производства. Проведенный анализ технологического процесса производства позволил выявить ключевые этапы и особенности функционирования системы. Разработанная система обладает значительным потенциалом для повышения эффективности производства, сокращения времени разработки и персонализации продукции. Она может быть использована в различных отраслях, таких как текстильная промышленность, производство упаковки, рекламная индустрия и др. Данная разработка является свидетельством растущего влияния искусственного интеллекта на современную промышленность и открывает новые возможности для создания инновационных решений.

## **Список используемых источников**

1.М.В. Ванчаков, И.С. Артамонов «ТЕХНОЛОГИЯ И ОБОРУДОВАНИЕ ДЛЯ ПРОИЗВОДСТВА КАРТОННОЙ И БУМАЖНОЙ ТАРЫ. Часть 2. Учебное пособие». Санкт-Петербург. 2022 г. Стр. 82-83 —

URL: <https://nizrp.narod.ru/metod/kafmavsystem/1673832710.pdf> (Дата обращения: 28.02.2024).

2. Казанский национальный исследовательский технологический университет. Дисциплина «Функциональные наноструктурированные полимерные материалы». Лекция №6. Стр. 2-3 — URL: [https://moodle.kstu.ru/pluginfile.php/339184/mod\\_resource/content/1/ФНПМ%20лб.pdf](https://moodle.kstu.ru/pluginfile.php/339184/mod_resource/content/1/ФНПМ%20лб.pdf) (Дата обращения: 28.02.2024).

3. «Пакет. Производство полипропиленовых пакетов». «Требования к макетам» — URL: <https://www.packet.msk.ru/dizajn-upakovki/trebovaniya-k-maketam/> (Дата обращения: 06.04.2024).

4. Уфимский Государственный Нефтяной Технический Университет. «Конспект лекция по компьютерной графике» — URL: <https://studfile.net/preview/2180165/page:45/> (Дата обращения: 06.04.2024).

5. Sg-company.ru. «Какое разрешение изображения следует использовать для этикеток?» — URL: [https://sg-company.ru/article/razresheniye\\_izobrazheniya\\_etiketok.html](https://sg-company.ru/article/razresheniye_izobrazheniya_etiketok.html) (Дата обращения: 06.04.2024).

6. Nikuson. «Как работает DALL-E» от 13.09.2022 г. — URL: <https://habr.com/ru/companies/ruvds/articles/687508/> (Дата обращения: 09.03.2024).

7. PatientZero. «Как работает Stable Diffusion: объяснение в картинка» от 17.10.2022 г. — URL: <https://habr.com/ru/articles/693298/> (Дата обращения: 09.03.2024).